

# VU Research Portal

## Meta-Analytical Studies in Transport Economics

Brons, M.R.E.

2006

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### **citation for published version (APA)**

Brons, M. R. E. (2006). *Meta-Analytical Studies in Transport Economics: Methodology and Applications*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam]. Thela Thesis.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

# **Meta-Analytical Studies in Transport Economics**

Methodology and Applications

ISBN 9051708912

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **378** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

**Meta-Analytical Studies in Transport Economics**  
**Methodology and Applications**

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan  
de Vrije Universiteit Amsterdam,  
op gezag van de rector magnificus  
prof.dr. T. Sminia,  
in het openbaar te verdedigen  
ten overstaan van de promotiecommissie  
van de faculteit der Economische Wetenschappen en Bedrijfskunde  
op donderdag 18 mei 2006 om 15.45 uur  
in de aula van de universiteit,  
De Boelelaan 1105

door

Martijn Reinout Eelco Brons

geboren te Amsterdam

promotoren:     prof.dr. P. Nijkamp  
                      prof.dr. P. Rietveld

# Acknowledgements

This Ph.D. thesis contains the result of research done at the Tinbergen Institute and the Department of Spatial Economics of the Vrije Universiteit Amsterdam. This research is realized within the framework of the Incentive Programme: Environment & Economics, funded by the Netherlands Organization for Scientific Research (NWO). Of course, I would never have reached the point of finishing my dissertation without the help and support of others.

First of all, I would like to thank Piet Rietveld and Peter Nijkamp for giving me the opportunity to work at this department, first as a student assistant and later as a Ph.D. candidate. I am indebted to them and to Eric Pels for their excellent supervision, their guidance and the freedom they gave me in doing research. Thanks for reading and re-reading my work and for providing many useful comments and suggestions.

In more than one respect, these four years have felt like a journey. Having said that, the first part of the journey oftentimes felt more like a survival trip. The Tinbergen M.Phil. program exceeded my expectations, to put it mildly. Already in the second week, the Tinbergen night pass replaced my Albert Heijn bonus card as the most frequented card in my wallet. I guess it were the lectures in cooperative game theory that kept Carla, Ernesto, Miguel, Robert, Sebi, Wendy and myself in the 'game'. Either those, or the spiritual support that we received on a regular basis from various Belgian Trappist orders (Westmalle, Chimay et al.), together with Hugo, Jasper, Marcos, Razvan, Sunčica, Vali and many others.

Soon afterwards, the Tinbergen Institute moved to a new location where I shared a room with Ana, Ghebre, José, Romy, Rute and Yin-Yen. Although Oceanic and Antarctic people were sorely missed, I enjoyed the 'multi-continental' atmosphere in our room. Moreover, the entrance was strategically located between the 'coffee room' and the coffee machine, which proved rather convenient during those nightly sessions shortly before deadlines. Many dinners, parties, movie nights, concerts and other get-togethers made the

period at the Tinbergen Institute very enjoyable. Of course, I owe many thanks to the secretaries of the Tinbergen Institute, in particular Marian and Nora, who have been of great support during this period.

After my period at the Tinbergen Institute, I moved to the department of Spatial Economics. Here I got a desk at the 'MasterPoint', which at that time functioned *inter alia* as an institute to transform Tinbergen students into fully functioning department members. I soon learned that being a Ph.D. student at this department does not only entail learning how to do research but also how to skate, play golf, build rope bridges and rafts, canoe at sea and so on. Fortunately, it also involves an (almost) weekly cake eating ritual, which provides ample opportunity to regain those lost calories. Thanks to the various 'Feestcommissies' for providing all of the above in addition to many enjoyable cultural events.

During most of my stay at the MasterPoint, I shared the room with Carla, Gert-Jan, Laura, Maria, Mark and Simonetta. I very much enjoyed our many conversations about work, music, movies, mensa 'delicacies' and how to deal with the printer (sweet words or violence?). I've also had a lot a fun with the next generation, including Moshe, Roberto, Sandy, Alex and the new HP 4250 printer, which is definitely an improvement.

I want to thank my colleagues at the department for many useful discussions, comments and suggestions at Eureka Seminars and MasterPoint meetings, and to all those people I have met on conferences, workshops and seminars, who have commented on my research. I would also like to thank Henri de Groot and Raymond Florax for their interest in working on joint papers. Henri's enthusiasm and his ability to make others enthusiastic about doing research have motivated me a lot early on. Raymond's expertise in meta-analysis has been of great help to me during these years. Of course, I should also thank Elfie. If not for her, the department would surely collapse and academic production would come to a standstill.

Many thanks also to the members of my promotion committee, David Banister, Jeroen van den Bergh, Raymond Florax, Rob van der Heijden en Bert van Wee, for spending time on reading my thesis and for providing me with many useful comments and suggestions.

Of course, there have been many people outside of academic life whom I want to thank for their continuous efforts to lead me astray from my academic obligations with many get-togethers and lengthy vacations. With respect to the latter, I want to thank 'the usual suspects', Camiel, Eddie, Eelco and Rob, but also Hugo and Maarten for having been fun travel companions on numerous journeys during these four years. Despite frozen toes in the Andes, tick attacks in the jungle, paragliding holidays under the rain, lost passports, stolen

wallets, chicken brain soup, camel liver kebab and increasing fuel taxes in the aviation sector (see also Chapter 4) we sure had a lot of fun together. Thanks and love to my parents for their encouragement and unrelentless support and to my sister Cathelijne who (finally) married Darco last year. Of course, Laura deserves a special word of thanks for her moral support, scientific support, companionship and, most importantly, for her love. I hope I have been equally supportive and loving in these years.

Thanks,  
Martijn





# CONTENTS

## PART I INTRODUCTION

<b>1 INTRODUCTION AND RESEARCH METHOD .....</b>	<b>3</b>
1.1 Mobility trends and external costs .....	3
1.2 Issues in transport policy making; valuation and reduction of external costs .....	6
1.3 Meta-analysis as a research approach .....	9
1.4 Research objectives and outline .....	11

## PART II META-ANALYSIS

<b>2 META-ANALYSIS: A METHOD FOR RESEARCH SYNTHESIS.....</b>	<b>17</b>
2.1 Introduction.....	17
2.2 Historical development and application.....	17
2.3 Meta-analysis: methodology and strengths .....	19
2.4 Statistical research techniques in meta-analysis.....	20
2.4.1 Univariate combining of effect sizes.....	21
2.4.2 Meta-regression analysis .....	23
2.5 Potential drawbacks and methodological issues .....	25
2.6 Meta-analysis in economic research .....	27
2.7 Conclusion .....	29
<b>3 ACCOUNTING FOR MULTIPLE SAMPLING IN META-ANALYSIS. A MONTE CARLO STUDY OF ESTIMATOR PERFORMANCE .....</b>	<b>31</b>
3.1 Introduction.....	31
3.2 Dealing with within-study dependence in meta-analysis .....	32
3.3 Experimental design.....	35
3.3.1 Generating dependence .....	35
3.3.2 Variation in design factors .....	36
3.3.3 The data generating process in steps .....	36
3.4 Estimation methods and evaluation criteria .....	37
3.4.1 Estimation methods.....	37
3.4.2 Evaluation criteria.....	39
3.5 Simulation results.....	40
3.5.1 Estimator performance.....	41
3.5.2 The effect of design factors on estimator performance .....	42
3.5.3 Response surface.....	45
3.6 Conclusion .....	49

# PART III APPLIED META-ANALYTICAL STUDIES IN TRANSPORTATION ECONOMICS

<b>4</b>	<b>A META-ANALYSIS OF PRICE ELASTICITIES OF DEMAND FOR AVIATION TRAVEL .....</b>	<b>53</b>
4.1	Introduction.....	53
4.2	Determinants of demand for passenger air transport.....	54
4.2.1	Choice contexts in air transport demand .....	54
4.2.2	Economic determinants of demand and price sensitivity .....	56
4.2.3	The effect of study characteristics on price elasticities .....	58
4.3	Meta-analysis and empirical results .....	59
4.3.1	Data collection and variables used .....	60
4.3.2	Descriptive statistics .....	61
4.3.3	Results of meta-regression analysis .....	65
4.4	Conclusion .....	68
	Appendix 4: List of primary studies on the price elasticity of aviation demand .....	70
<b>5</b>	<b>A META-ANALYSIS OF PRICE ELASTICITIES OF GASOLINE DEMAND. A SYSTEM OF EQUATIONS APPROACH.....</b>	<b>71</b>
5.1	Introduction.....	71
5.2	Price sensitivity in the context of gasoline demand .....	72
5.3	A model based on a system of meta-regression equations: the SMR model.....	76
5.4	Constrained estimation of mean effect sizes .....	78
5.5	Results of meta-regression analysis .....	80
5.6	Summary and conclusion .....	86
	Appendix 5: List of primary studies on price elasticities of gasoline demand .....	90
<b>6</b>	<b>A META-ANALYSIS OF TECHNICAL EFFICIENCY IN URBAN PUBLIC TRANSPORT .....</b>	<b>91</b>
6.1	Introduction.....	91
6.2	The concept of technical efficiency .....	92
6.3	The construction of efficiency frontiers.....	94
6.4	Study comparability issues.....	96
6.5	Quantitative overview of the literature on public transport efficiency.....	99
6.6	Results of meta-regression analysis .....	104
6.7	Summary and concluding remarks.....	107
	Appendix 6: List of primary studies on technical efficiency of urban public transport.....	109
<b>7</b>	<b>RAIL NOISE. ECONOMIC VALUATION AND POLICY.....</b>	<b>111</b>
7.1	Introduction.....	111
7.2	Rail noise, external costs and policy: a conceptual model .....	112
7.3	Government policy: emission standards .....	113
7.4	Noise emission and immission.....	114
7.4.1	Noise emission sources and reduction measures.....	114
7.4.2	Noise dispersion.....	115
7.5	Disturbance from rail noise.....	116
7.6	The economic valuation of rail noise .....	118
7.6.1	Indirect damage costs.....	120
7.6.2	Direct costs: property value as a proxy .....	120
7.6.3	Prevention costs .....	122
7.7	Conclusion .....	124
	Appendix 7 : Overview of primary studies on the economic valuation of rail noise.....	127

## **PART IV SYNTHESIS**

### **8 CONCLUSIONS AND DISCUSSION .....131**

8.1	Methodological insights .....	132
8.1	Empirical results .....	134
8.3	Policy implications.....	136
8.4	Directions for further research .....	140

### **REFERENCES.....143**

### **SAMENVATTING (SUMMARY IN DUTCH).....163**



# **PART I**

## **INTRODUCTION**



# Chapter 1

## Introduction and Research Method

### 1.1 Mobility Trends and External Costs

Two-thousand years ago, the streets of Rome had become fetid and knotted with traffic. Local rulers became so fed up that they declared that the circulation of the people should not be hindered by numerous litters and noisy chariots. It was an early salvo in what would become an endless, thankless, unwinnable war. Around the same time, Julius Caesar introduced the first off-street parking laws. In A.D. 125, a limit was placed on the number of vehicles that were allowed to enter Rome. For as long as there have been roads, it seems, there have been crowds of swearing, sweaty drivers — and schemes to get rid of them (Ripley, 2003).

Vast improvements in living standards and transportation systems in the twentieth century have instigated a spectacular growth in the demand for motorized transportation worldwide. Statistics indicate a structural rise in traveler mobility, while the increasing liberalization of trade has stimulated a drastic increase in trans-boundary freight transport (see Reggiani et al., 1999). Recent studies by *inter alia* the European Environment Agency (EEA) and the Organization for Economic Co-operation and Development (OECD) show a number of disquieting trends. Transport volumes and the number of motor vehicles in Europe have been growing steadily over the past 30 years. In the European Union, passenger and freight transport have more than doubled during the past 25 years and car ownership is approaching one car for every two inhabitants (EEA, 2000a, b). The pace of this growth follows that of GDP.

Road transport is continuing to increase its market share compared to other modes. In developed market economies, this is true for passenger and goods transport alike. Private passenger cars now account for more than 80% of traffic volumes there. It applies also to countries in transition, where individual passenger transport is widely seen as an expression of personal freedom and economic success and where goods transport, due to a lack of competitive alternatives, is increasingly dominated by road transport. Public and rail transport used to play an important role in Europe but they are quickly losing ground to private road

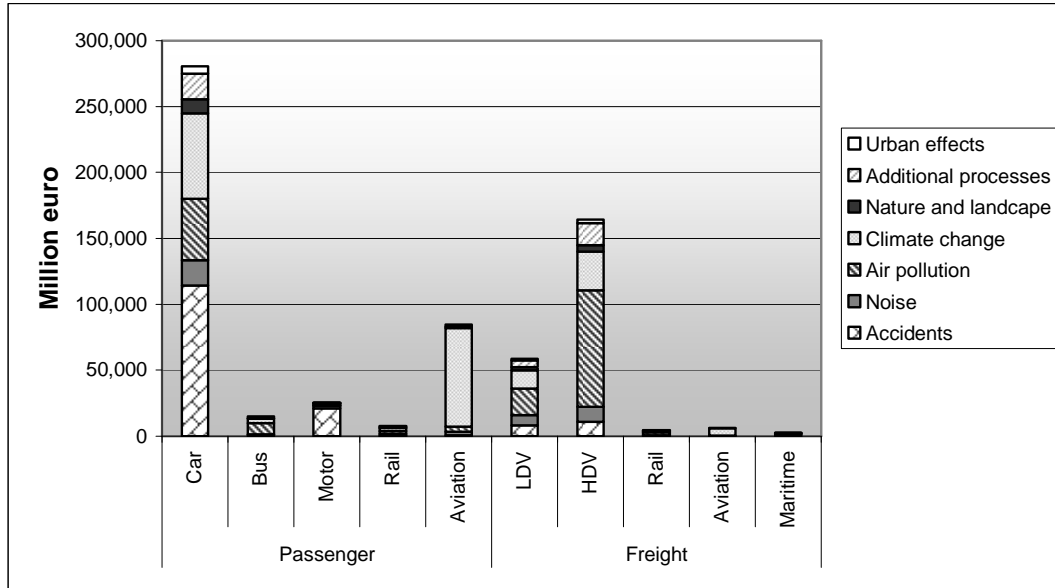


transport, in part due to a lack of investment and maintenance of infrastructure and fleets. A study of 14 Central- and East-European countries and newly independent states predicted that, if current policies continue, by 2010 passenger car use will have doubled compared to 1994 levels; by 2030, it will have increased a further 150%. Road freight traffic is expected to increase even more rapidly (CEI/UNEP/OECD/Austrian Federal Ministry for Environment, Youth and Family, 1999). Aviation is the fastest growing mode of passenger transport, at first due to income growth and later because of the entry of low-cost carriers on the aviation market. According to IATA, European passenger air traffic more than doubled between 1985 and 1998 (an average growth of almost 7% a year) and is expected to continue to grow. Between 1998 and 2015 it is estimated that European passenger air traffic will also more than double - to about 1,100 million passengers a year (IATA, 2000). Maritime transport is likewise increasing. During the past decade, there has been an increase of approximately 5% a year on a global level (SAI, 2001).

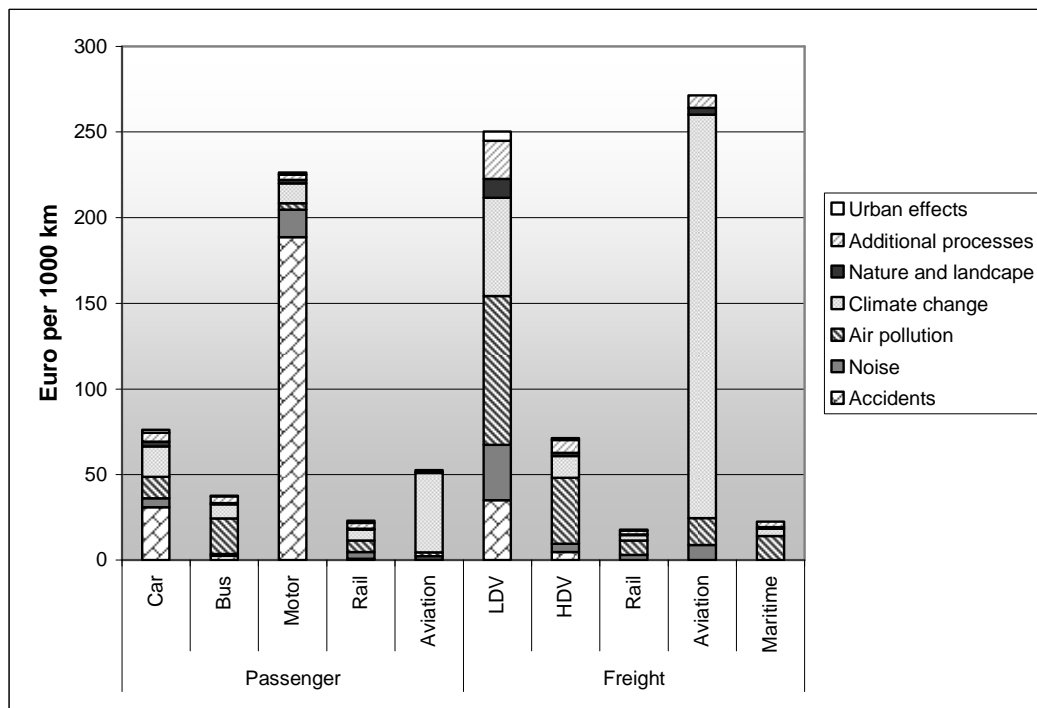
The availability of efficient and effective transportation networks is a crucial precondition for economic development and an asset for local, regional and international accessibility. Historical evidence indicates a strong correlation between economic growth and growth in transport. However, the explosive growth in transport has also caused an increase in adverse external effects. Road congestion in many urban areas has been increasing in duration and intensity. Transport produces more air pollution than any other human activity. Traffic noise affects to a serious extent almost half the urban residents in most countries. Road safety is a major concern in cities and elsewhere. Carbon dioxide (CO<sub>2</sub>), which is emitted whenever fossil fuels are burnt, spreads to the upper atmosphere and contributes to global warming. Increased car dependence results in traffic domination in urban areas and to increasing isolation for those without access to a car. This result is exacerbated by the fact that fringe areas of cities are difficult and costly to serve by public transport. Finally, transport contributes to the decaying urban fabric and neglect of central city areas, as well as urban sprawl (OECD/ECMT, 1995).

In the EUR 17 countries, total external costs (excluding congestion costs) rose by more than 12 per cent in the period 1995-2000 towards roughly 7.3 per cent of the GDP (Infras/IWW, 2004). Figure 1.1 shows the total external costs per transport mode in 2000. Road transport is responsible for the vast majority (83.7 per cent) of the total external transport costs, followed by the aviation sector (14 per cent). The costs related to rail transport and interior navigation are significantly lower (1.9 per cent and 0.4 per cent of total external costs, respectively). Two-third of the external costs is caused by passenger transport and the

remaining third by freight transport. Figure 1.1 furthermore indicates that in the case of road transport most of the external costs are caused by accidents, air pollution and climate change. In the aviation sector, climate change is responsible for the majority of the external costs.



**Figure 1.1:** Total external costs per transport mode for EUR 17. LDV = light duty vehicle. HDV = heavy-duty vehicle. Source: Infrac/IWW (2004).



**Figure 1.2:** Average external costs per transport mode for EUR 17. LDV = light duty vehicle. HDV = heavy-duty vehicle. Costs are shown per 1000 passenger kilometers for passenger transport and per 1000 ton kilometers for freight transport. Source: Infrac / IWW (2004).

Figure 1.2 shows the average external costs per transport mode. In the passenger transport sector, the external costs per passenger kilometer are considerably lower for train and bus than for car and airplane. Given the high value of the total external costs of automobile transport and aviation transport, this indicates that the total external costs could be substantially reduced if automobilists switch to bus and rail transport. Also in the freight transport sector, the external costs per ton kilometer are the highest for road transport and aviation, while rail and maritime transport generate the lowest average costs. As is the case in the passenger transport sector, intermodal substitution can lead to a large reduction in total external costs.

## **1.2 Issues in Transport Policy Making; Valuation and Reduction of External Costs**

The explosive growth in external costs in the late twentieth century has caused national as well as international governmental bodies to worry about the sustainability of their transport systems (see Banister, 1999; Tolley, 2004; Newman and Kenworthy, 2004). A sustainable transportation system: (i) allows the basic access needs of individuals and societies to be met safely and in a manner consistent with human and ecosystem health, and with equity within and between generations; (ii) is affordable, operates efficiently, offers choice of transport mode, and supports a vibrant economy; (iii) limits emissions and waste within the planet's ability to absorb them, minimizes consumption of non-renewable resources, limits consumption of renewable resources to the sustainable yield level, reuses and recycles its components, and minimizes the use of land and the production of noise (Centre for Sustainable Transportation, 2002)<sup>1</sup>. Sustainable transport policy should aim at developing and supporting such a transportation system.

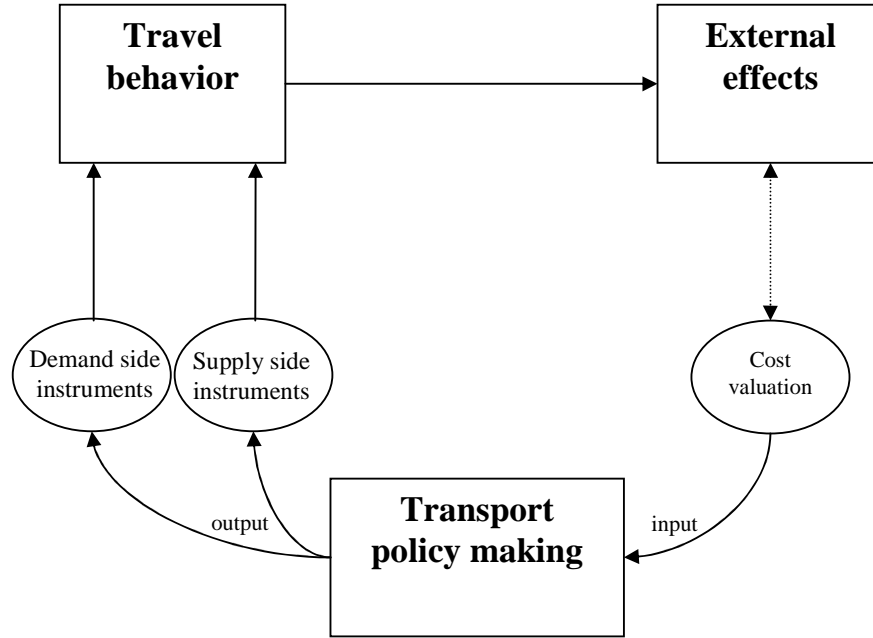
Ensuring that transport policy serves the goals of sustainable development requires a balance between striving for economic goals and the social and environmental pillars of sustainable development. Exploring how economic growth can be achieved without stimulating transport growth is therefore a prerequisite to ensuring that the economic, social and environmental goals for transport policy can be mutually supportive (EU, 2001). The conclusions from the European Council in Gothenburg in June 2001 state that a sustainable transport policy should tackle rising volumes of traffic and levels of congestion, noise and pollution and that action is needed to bring about a significant decoupling of transport growth

---

<sup>1</sup> This definition of sustainable transport is developed by The Centre for Sustainable Transportation in Canada and has also been adopted by the European Council of Ministers for Transport.

and GDP growth, in particular by a shift from road to rail, water and public passenger transport." The volume of transport is also a concern in the 6th Environmental Action Programme, which calls for decoupling of economic growth and the demand for transport with the aim of reducing environmental impacts. Yet the extent to which transport growth can be decoupled from economic growth remains disputed. Transport growth persists and is strongly linked to economic growth; many commentators view this as inevitable. Although governments have not deliberately attempted to decouple transport and economic growth, the belief persists that there is not just a link, but a causal relationship. Moreover, evidence to support decoupling remains limited and somewhat theoretical (EU, 2001).

Figure 1.3 shows that environmental policy making in transport can be described by a feedback process. Demand for transportation causes external effects to occur. These effects are expressed in monetary terms by valuation techniques. Externality valuation provides major contributions to the formulation of sustainable development policies (see Ricci, 2004). On the one hand, supply side policies (infrastructure, products and services in general) require that the relative merits of alternative options be assessed and compared on an equitable basis. These policies can therefore directly benefit from: (i) the availability of an extended accounting framework for cost benefit analyses, where all social costs (i.e., both internal and external) are considered in a consistent and homogeneous manner, and (ii) the subsequent possibility of comparing alternative investment options (e.g. through the assessment of differential values). On the other hand, demand side policies (regulation, pricing and taxation) should systematically avail themselves of an accurate estimation of the absolute value of externalities in order to define the optimal terms and conditions of mitigation/abatement measures and measures for the internalization of external costs. This need for external cost valuation has given rise to a large number of valuation studies. Valuation studies often aim to establish an index function that expresses the monetary costs of the external effect per unit. Due to heterogeneity in valuation methods and other case-specific characteristics, the validity of such an index function is constrained by the characteristics of the study in question; the results of the study are 'context-specific'. The standard argument is that the more results depend on the context in which they are obtained, the less likely it is that they can be applied to a new, similar kind of study conducted in the future (Bal et al., 2002). Of course, this also holds for the application of previously calculated valuation indices to new case studies that may have rather different characteristics.



**Figure 1.3:** The role of transport policy making in the reduction of the external costs of transport.

Demand side sustainable transport policy tries to influence the behavior of transport users in order to minimize the costs imposed to society as a whole. The most important market-based instruments used in this context are pricing instruments. In practice, these are based on fiscal and other taxation mechanisms that are expected to produce the necessary price adjustments. Pricing policy often focuses on the reduction of demand for a specific transport mode or the redistribution of demand over different transport modes. Empirical knowledge of the degree of price sensitivity in the transport sector is essential information in determining the price level that is required to achieve the desired reduction or redistribution of demand. Nijkamp and Pepping (1998) mention the price elasticity of transport demand as the most important parameter to understand how pricing policies will affect transport demand. With respect to transport demand, price sensitivity is dependent on many factors including the type of price change, the type of trip and traveler, the availability of substitutes, the scale and scope of pricing policy and the time period that is studied (VTPI, 2002). Hence, the practical value of any empirical estimate of price elasticities of demand for transport is constrained by the characteristics of the empirical study in question. This decreases the accuracy of cost-benefit analyses that are based on previously elasticity estimates.

Supply side sustainable transport policy includes the provision of transport infrastructure and non-infrastructure goods and services. An efficient and well operating

public transport system is an essential component of a sustainable transportation system. In combination with well-directed pricing policy, such a public transportation system is a potentially powerful policy instrument to reduce external costs by means of a reduction or redistribution of transportation demand. Based on market limitations there are sound economic reasons for a significant degree of state intervention in the field of public transport. It has become an important role of the state in transportation matters to secure and provide all the country's citizens with urban transport service. With the aim to carry out this goal, it is often necessary to heavily subsidize the urban transportation system. In the past two decades, serious concerns about possible regulatory failures have led to a reassessment of the role of the state in the organization of the urban transportation sector (see Glaister et al., 1990). In view of these concerns, it is of great interest to investigate whether urban transport operators work in a technically efficient way (i.e., reach economic targets such as cost minimization or output maximization conditional on output or input constraints). Solid technical efficiency measurement can provide a basis for a thorough discussion on the relative merits of private versus public provision of transportation services. A comparative research approach is particularly appealing as it allows for the comparison of the technical efficiency level estimated by various case studies with different characteristics. Hence, such an approach may identify interesting links between the level of technical efficiency and case-specific characteristics.

### **1.3 Meta-Analysis as a Research Approach**

The availability of numerous empirical assessment studies on the previously discussed key-issues of environmental transport policy contributes to the relevance of the pursuit of synthetic views on these issues. Hence, a rigorous statistical analysis in a comparative research framework is an appealing and useful approach. *Meta-analysis* is such an approach. Meta-analysis refers to the statistical summary of previously documented study results. It is, in effect, a study of studies, aimed at a higher level of analysis than that produced by the limited viewpoint of a single investigation. Modern meta-analysis was initiated by the work of Glass (1976) on psychotherapy and spread rapidly in the social and experimental sciences such as education and psychology, and especially the health sciences. More recently, meta-analysis has become a well-accepted tool in micro-economically oriented disciplines such as environmental and natural resource economics (for an overview see Florax, 2002c). In the

field of transportation economics, various interesting meta-analytical studies have been conducted (see Button, 1995).

Meta-analysis is based on the collection of previously documented study results with respect to a certain effect of interest, together with relevant quantitative information that describes the characteristics of the primary studies. Data collection is often based on *multiple sampling* techniques, i.e. multiple observations are collected per primary study. This is particularly common in meta-analyses in the field of economics. Next, statistical techniques are used in order to derive meaningful insights with respect to the variable of interest and information about the relationship between the variable of interest and the study characteristics. These statistical techniques include descriptive statistics, analysis of variance, vote counting, combining of effect sizes and regression analysis techniques. As such, the quantitative character of meta-analysis provides a rigorous complement to the casual, narrative discussions of research that typify the conventional literature reviewing studies. An important asset of the meta-analytical method lies in the fact that, by pooling study results together, more reliable and precise estimates of the effect size can be obtained that are not constrained by the characteristics of a single empirical study. Furthermore, *meta-regression analysis* enables the analyst to estimate the impact value and significance of conditioning factors on the effect size. These results can be used to estimate the “expected” value of the variable of interest in a specific setting.

The added value of meta-analysis for the valuation of transport externalities has previously been illustrated by meta-analytical studies on congestion policy (Button and Kerr, 1996), the valuation of aircraft noise (Schipper et al., 2001; Nelson, 2004), the value of statistical life in road safety (de Blaeij et al., 2003) and external transport cost estimates (Quinet, 2004). By combining previously documented estimates of an external cost index, while controlling for conditional factors, conditional mean estimates can be produced. The scope for practical application of such a conditional estimate is wider than that of primary study estimates in the sense that the mean estimate is less context-specific and has increased statistical power. Meta-analysis furthermore enables the analyst to identify the impact of determinant factors on the external cost index. Conditional means, in combination with knowledge on the impact of conditioning variables, may be used for the application of value transfer and conditional forecasting. As such, meta-analysis can facilitate the development of uniform structures for valuation and assessment. This is a major asset for sustainable transport policy.

For the analysis of price sensitivity, the advantages of a meta-analytical approach are similar, as demonstrated by studies on public transport demand elasticities (Nijkamp and Pepping, 1998), price and income elasticities of gasoline demand for gasoline (Espey, 1998) and price elasticities of transport demand in general (Kremers et al., 2000). Conditional estimates can be obtained that have a wider range of applicability than estimates produced by primary studies. Identification of the impact of determinant factors of the price elasticity can provide meaningful insight into the price behavior of consumers as well as valuable information about the “expected” degree of price sensitivity in a specific setting. Such information facilitates the development of well-directed pricing policy.

Also with respect to the analysis of the technical efficiency of public transport, meta-analysis can be used to estimate mean conditional values and to identify the impact of determinant factors. If determinant factors are under the control of authorities, they can be used as policy instruments in order to increase the overall technical efficiency and therefore increase the competitive strength of urban public transport. More specifically, meta-regression analysis can provide important insights into the relative merits of private versus public provision of transportation services. Competitive public transport in combination with well-directed pricing policy is a potentially powerful combination of environmental policy instruments.

## **1.4 Research Objectives and Outline**

In this chapter, we discussed a number of key-issues in environmental transport policymaking, i.e. the effectiveness of pricing policy in reducing demand for transport, the efficiency of urban public transport and the monetary valuation of transportation externalities. In the framework of this thesis, we aim to obtain greater insight into these key-issues by means of a set of meta-analytical studies in which we focus on indicators of price sensitivity of transport demand, efficiency of public transport and external costs of transport. In each of these studies, we investigate the indicator of interest in the context of a specific transport mode. More specifically, we aim to apply existing and newly developed meta-analytical techniques in order to explain the empirical variation in (i) the price elasticity of demand for aviation transport, (ii) the price elasticity of demand for gasoline demand, (iii) the technical efficiency of urban public transport and (iv) the economic costs of rail noise per noise unit.

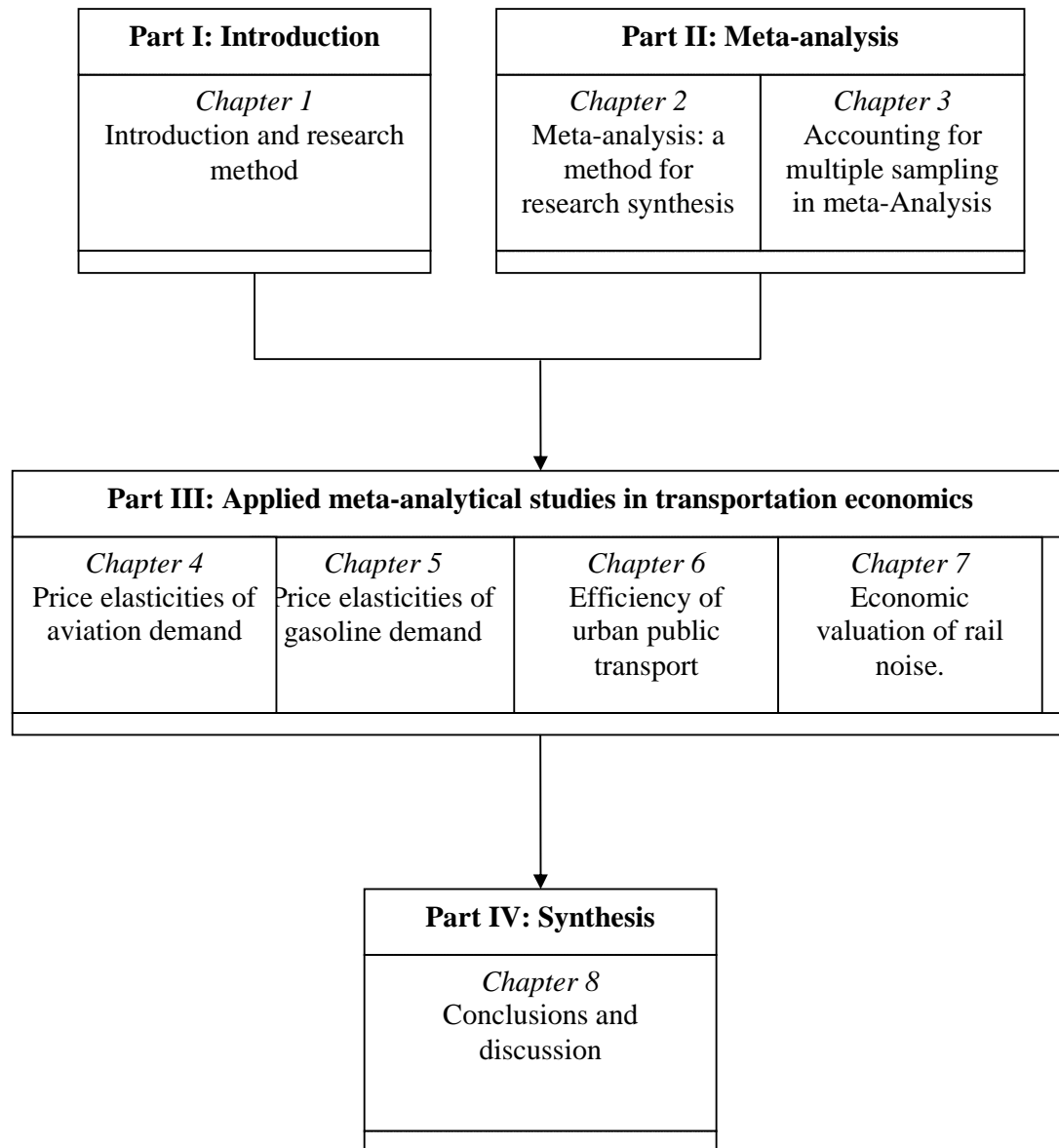


In doing so, we focus on the estimation of combined mean values for these indicators and on identification of the impact of study characteristics and other conditioning factors on the value of these indicators. Furthermore, in order to ensure the use of appropriate estimation techniques, we compare the performance of various estimation techniques in accounting for statistical within-study dependence, which is a common problem in meta-analysis datasets in economics.

The outline of this thesis is shown in Figure 1.4. In Part II of this thesis, we focus on the methodological framework that underlies this dissertation, i.e., meta-analysis. First, in Chapter 2 we give a brief description of the history, background, strengths and methodology of meta-analysis as a research method. Furthermore, in this chapter we discuss a number of potential drawbacks of meta-analysis. Next, in Chapter 3 we present a Monte Carlo simulation study in which we focus in more detail on one of these drawbacks, i.e. within-study dependence associated with multiple sampling techniques. We compare the performance of several estimation techniques for doing meta-regression analysis on datasets that are based on multiple sampling. The results of this study provide us with important methodological insights that we use in the applied work in Part III of this thesis.

Part III consists of four studies, described in Chapters 4, 5, 6 and 7, which form the core of this thesis. Each of these studies investigates one of the three key-issues of sustainable transport policy in the context of a specific research topic. In Chapter 4, we present a meta-analytical study that investigates the variation in the price elasticities of demand for passenger air transport. In this study, we use meta-regression analysis in order to estimate the impact of study characteristics and other conditioning factors on the price elasticity. The meta-analytical study presented in Chapter 5 uses a similar approach but focuses on price elasticities of gasoline demand. In this chapter, we develop a new meta-analytical method based on a system of meta-regression equations that enables us to use more underlying information and to derive more detailed results than a conventional meta-analytical approach. We use this method to estimate mean values for a set of related price elasticities and to analyze the impact of study characteristics and conditioning factors on the values of the price elasticities. In the meta-analytical study described in Chapter 6, we investigate the variation in the technical efficiency of urban public transport operators. Again, we use meta-regression analysis to investigate the impact of various study characteristics and other conditioning factors on the level of technical efficiency. Chapter 7 presents a review study, which investigates the economic impact of noise from rail transport. In Chapter 8, we summarize our findings and

we provide methodological and empirical conclusions as well as directions for further research.



**Figure 1.4:** Thesis outline



## **PART II**

### **META-ANALYSIS**



# Chapter 2

## Meta-Analysis: a Method for Research Synthesis

### 2.1 Introduction

The essential character of meta-analysis is the statistical analysis of the findings of many empirical studies (Glass, 1976). While the results of most (social science) studies are constrained by the populations from which subjects are drawn, by the limitations of their measures, and by various additional methodological constraints, the meta-analytical method offers the capacity to produce more reliable and precise results that are based on the findings from a series of studies. It is, in effect, a study of studies, aimed at a higher level of analysis than that produced by the limited viewpoint of a single investigation.

In the present chapter, we describe the background and methodology of meta-analysis as a research method. First, in Section 2.2 we briefly describe the historical development of meta-analysis. Furthermore, we provide a brief literature overview. In Section 2.3, we discuss the methodology and assets of meta-analysis as a research method. In Section 2.4, we discuss a number of univariate and multivariate techniques that are commonly used in meta-analysis. In Section 2.5, we focus on a number of drawbacks and methodological issues in meta-analysis. Section 2.6 discusses some issues that are specific to meta-analysis of economic research. Section 2.7 concludes.

### 2.2 Historical Development and Application

Although the term *meta-analysis* was first coined by Glass in 1976, the problem of combining the results of quantitative research has a much longer history. Meta-analyses that involve some form of averaging estimates from different studies have been part of the statistical literature since the beginning of the 20th century. However, the modern era of meta-analysis began with the work of Glass on psychotherapy (Glass, 1976; Smith and Glass, 1977; Smith et al., 1980), Schmidt and Hunter (1977) on validity coefficients for employment tests, and Rosenthal and Rubin (1978) on interpersonal expectancy effects. Partly because of the major

impact of the work of Glass and his group, later years have shown a rapidly growing number of investigators who have been discussing, employing and developing a variety of meta-analytic procedures. Meta-analysis spread rapidly in the social and experimental sciences such as education and psychology, and especially in the health sciences where it has been virtually institutionalized as the preferred approach to integrating the findings of clinical trials research (see for example Olkin, 1992 and Sacks et al., 1987).

Meanwhile also the methodological and statistical work has been expanded and sophisticated. Hedges and Olkin (1985) provided the rigorous statistical proofs that established meta-analysis as an independent specialty within the statistical sciences. Cooper and Hedges (1994) and, more recently, Sutton et al. (2000) provide good overviews of methodological techniques applied in meta-analysis.

Meta-analysis is by now a well-accepted tool in macro-economically oriented fields such as labor economics and in more micro-economically oriented fields such as environmental and natural resource economics. Examples of macro-economic oriented meta-analyses include studies on labor participation and productivity (Doucouliagos, 1995), minimum wage and unemployment (Card and Kruger, 1995), wage elasticities of labor demand (Espey and Thilmany, 2000), productivity spillovers of multi-national firms (Görg and Strobl, 2001), union and non-union wage differentials (Jarrell and Stanley, 1990), environmental regulation and business location (Jeppesen et al., 2001), growth effects of fiscal policy (Philips and Goss, 1985; Nijkamp and Poot, 2002) and Ricardian equivalence (Stanley 1998). In the field of environmental and natural resource economics, many meta-analyses focus on valuation-related research. Meta-analyses within this field include studies on the valuation of air pollution (Smith and Huang, 1995; Schwartz, 1994), the valuation of biodiversity (Loomis and White, 1996; Espey and Kaufman, 2000), price and income elasticities of residential water demand (Dalhuisen et al., 2001; Espey et al., 1997), the valuation of wetland services (Woodward and Wui, 2001; Brander et al., 2003; Brouwer et al., 1999), the effect of environmental regulation on competitiveness (Mulatu et al., 2002) and valuation methodology (List and Gallet, 2001; Smith and Osborne, 1996; Smith and Pattanayak, 2002).

Also in the field of transportation economics, various meta-analytical studies have been conducted. Examples of research topics that have been investigated in a meta-analysis framework are public transport demand elasticities (Nijkamp and Pepping, 1998), congestion policy (Button and Kerr, 1996), the valuation of aircraft noise (Schipper et al., 2001; Nelson, 2004), the value of statistical life in road safety (de Blaeij et al., 2003), external transport cost

estimates (Quinet, 2004), elasticities of gasoline demand for gasoline (Espey, 1998), road marking policy (Davidse et al., 2004) and travel time valuation (Wardman et al., 2003). In Button (1995), the practical use of meta-analysis for transport economic research is demonstrated by applied meta-analyses on travel time savings, traffic noise valuation, and land use impact of transport.

## 2.3 Meta-Analysis: Methodology and Strengths

There are many areas in social science for which results of a large number of studies are available that all address essentially the same question. Qualitative summarizations of the results of these studies, while certainly useful, are unable to use all the potential information that the set of studies has to offer, either with respect to summarized effect sizes or with respect to summarized significance levels. Often, they do not tell us more about each study in a set of studies than the direction of the relationship between the variables investigated and whether or not a given significance level was attained. This state of affairs is changing. Glass (1976) promotes meta-analysis as a rigorous alternative to the casual, narrative discussions of research, which typify our attempt to make sense of the rapidly expanding research literature. Glass (1976) defines meta-analysis as follows: “meta-analysis refers to the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings”. Meta-analysis is, in simple terms, the use of formal statistical techniques to sum up a body of separate but similar studies. It provides a series of techniques that allow the cumulative results of a set of individual studies to be pulled together (Button, 1995). The study results that are to be combined and investigated are collected in the form of estimates of a specific indicator of interest, which, in the literature on meta-analysis, is called *effect size*.

The main advantage of meta-analysis is that, by pooling all the study results together, one can reduce the effect of random error and so produce more reliable and precise estimates of the effect size of interest that are not constrained by the characteristics of a single empirical study. Furthermore, meta-analysis can be used to investigate the heterogeneity among studies. This enables the analyst to identify the impact of study characteristics and other conditioning factors on the value of the effect size. These applications of meta-analysis are described in more detail in Sections 2.4 and 2.5.

In the experimental sciences, the effect size is usually a measure of the relationship between any two variables that is based on a correlation coefficient or on the standardized



mean differences between experimental and control group (Keef and Roberts, 2004). In meta-analysis in the economic literature, the effect size typically has the form of an elasticity or a nominal value, such as consumer surplus or willingness to pay.

A meta-analytical approach is particularly attractive in cases where a *sufficient number of studies* report *comparable* research results in a *quantitative* format. Indeed, more and more reviews of the literature are moving from the traditional qualitative literary format to the quantitative format (Rosenthal, 1991). When comparing meta-analysis with qualitative research reviewing techniques, four differences constitute the primary advantages of meta-analysis (based on Lipsey and Wilson, 2001). First, good meta-analysis is conducted as a formal and structured research technique in its own right. Definition of the study population, study retrieval, coding of study characteristics and data analysis are done in an organized, systematic and explicit way while conclusions are supported by the results of the data-analysis. Second, meta-analysis represents key study findings in a fashion that is more differentiated and sophisticated than conventional review procedures in the sense that its effect sizes constitute a variable sensitive to findings of different strength across studies. Third, the systematic coding of study characteristics in meta-analysis permits an analytically precise examination of the relationships between study findings and study characteristics that is not available in other approaches to summarizing research. Finally, meta-analysis provides an organized way of handling information from a large number of study findings under review.

## 2.4 Statistical Research Techniques in Meta-Analysis

Meta-analysis encompasses a broad selection of statistical research synthesis methods. Commonly applied techniques include descriptive analysis of effect sizes, analysis of variance, vote counting, combining of effect sizes and meta-regression analysis.

From a descriptive analysis of the sample of effect sizes, supported by distributional graphs and descriptive statistics, the analyst may already gain some meaningful insights with respect to the object of study. Analysis of variance (ANOVA) is used to determine differences among the mean effect sizes of two or more groups of observations. The observations can be grouped or categorized according to certain study characteristics, such as study methods, geographical setting or functional form. Vote counting is a simple method that can be used when the primary studies do not contain enough information to calculate effect size

estimates<sup>2</sup>. The effect sizes are grouped into three categories, i.e., significantly positive, significantly negative and insignificant. The number of effect sizes in each of the categories is counted. The category that has the most estimates is assumed to give the best indication of the direction of the true relationship (Cooper and Hedges, 1994).

A large number of meta-analyses focus on the univariate or multivariate combining of effect sizes in order to calculate mean effect size estimates. The literature distinguishes four main models that are commonly used for the combining of effect sizes (see for example Sutton et al., 2000 or Lipsey and Wilson, 2001): the fixed effects model, the random effects model, the fixed effects regression model and the mixed effects model<sup>3</sup>. The model of choice depends on assumptions or knowledge about the structure of the variation between the effect sizes.

### 2.4.1 Univariate Combining of Effect Sizes

The fixed effects model for combining effect sizes assumes that there is no heterogeneity between the study results; the studies are assumed to be estimating a single true underlying effect size. This implies that there is no variation between effect sizes in addition to sampling error. The mean effect size is computed by weighting each effect size by the inverse of its variance. For  $i = 1..n$  independent studies to be combined, let  $y_i$  be the observed effect size and  $v_i$  the variance of the underlying population effect size, for the  $i$ th study. The general formula for the weighted mean effect size is:

$$\bar{y}_F = \frac{\sum w_i y_i}{\sum w_i} \quad (2.1)$$

With  $w_i = 1/v_i$  and  $v_i = \text{var}(y_i)$ . The confidence interval for a mean effect size is based on the standard error of the mean and the  $z$ -value corresponding to the confidence interval. The standard error of the mean is computed as the square root of the sum of the inverse variance weights (Hedges and Olkin, 1985) as:

---

<sup>2</sup> Often the information on the effect size is in the form of a report of the decision yielded by the significance test (e.g., significant positive relation) or in the form of a direction of the effect without regard to its statistical significance (e.g., a positive mean difference or a positive correlation) (Cooper and Hedges, 1994).

<sup>3</sup> These models all weight for the inverse of the variance models, which is common practice in meta-analysis. Unweighted variants of these models such as the unweighted meta-regression model are sometimes used in the literature, for example in cases where the standard errors are unobserved or when they are assumed to be equal.

$$se(\bar{y}_F) = \sqrt{\frac{1}{\sum w_i}} \quad (2.2)$$

If the distribution of effect sizes around their mean is such that it can be explained from sampling error alone the use of the fixed effects model is appropriate. In order to test for this, a homogeneity test can be used that is based on the  $Q$ -statistic, which is distributed as a chi-square with  $n-1$  degrees of freedom where  $n$  is the number of effect sizes. The formula for the  $Q$  statistic is:

$$Q = \sum w_i (y_i - \bar{y}_F)^2 \quad (2.3)$$

If  $Q$  exceeds the critical value for a chi-square with  $n-1$  degrees of freedom, then the null-hypothesis of homogeneity is rejected. This indicates a heterogeneous distribution; the distribution of effect sizes is greater than what can be explained from sampling error and a different model should be adopted. The random effects model, the fixed effects regression model and the mixed effects model are all based on the assumption of a heterogeneous distribution of the effect size.

The random effects model assumes that each observed effect size differs from the population mean by observation-level sampling error *plus* a value that represents other sources of variability that are assumed to be randomly distributed. The sum of these two variance components,  $v_i^*$ , is the total variance associated with the distribution of effect size values and is described by  $v_i^* = v_\theta + v_i$  where  $v_\theta$  is the estimate of the random or between-studies variance component and  $v_i$  is the estimate of the variance associated with sampling error. The effect sizes are combined in a similar way as in the fixed effects model:

$$\bar{y}_R = \frac{\sum w_i^* y_i}{\sum w_i^*} \quad (2.4)$$

However, we have that  $w_i^* = 1/v_i^*$ . The standard error of the mean is computed as the square root of the sum of the inverse variance weights as:

$$se(\bar{y}_R) = \sqrt{\frac{1}{\sum w_i^*}} \quad (2.5)$$

### 2.4.2 Meta-Regression Analysis

If some of the variation beyond sampling error is systematic in the sense that it can be related to identifiable variation in study characteristics, a multivariate model should be adopted. The *fixed effects regression model* and the *mixed effects model* are multivariate models that enable the analyst to investigate and correct for the systematic variation in the effect size that is caused by the variation in study characteristics and other conditioning variables:

The fixed effects regression model assumes that all variation in the effect sizes beyond sampling error is systematic. This systematic variation is accounted for by including a set of moderator variables that capture the effect of study characteristics and other conditioning variables on the effect size. Suppose, as before, that there are  $n$  independent effect size estimates  $y_1, \dots, y_n$ , with estimated sampling variances  $v_1, \dots, v_n$ . Suppose also that there are  $k$  known moderator variables  $\mathbf{x}_1, \dots, \mathbf{x}_k$  which are believed to be related to the effects via a linear model of the form:

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \mu_i \quad (2.6)$$

where  $x_{i1}, \dots, x_{ik}$  are the values of the moderator variables  $\mathbf{x}_1, \dots, \mathbf{x}_k$  for the  $i$ th study, and  $\beta_0, \beta_1, \dots, \beta_k$  are the regression coefficients to be estimated. The coefficients, which represent the impact of the moderator variables on the effect size, are calculated via weighted least squares algorithms, with the weights defined by the inverse of the estimated sampling variances. Thus  $w_i = 1/v_i$  with  $v_i = \text{var}(\mu_i)$ . If, after removing the systematic variation that is associated with the variation in the moderator variables, the remaining variation can be explained from sampling error the use of the fixed effects regression model is appropriate. In order to check for remaining heterogeneity, a test can be used that is based on the  $Q$ -statistic of the weighted sum of squared residuals:

$$Q_E = \sum w_i [y_i - (\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik})]^2 \quad (2.7)$$

If  $Q_E$  exceeds the critical value for a chi-square with  $n - k - 1$ , then variation beyond sampling error remains across the effect sizes. This indicates that, even after removing the systematic variation, the effect sizes remain heterogeneous. In this case, a mixed effects model might be a better model of choice. The mixed effects model assumes that the variation beyond sampling error has both systematic and random sources of variation. As in the fixed effects model the systematic variation is accounted for by a set of moderator variables. As in the random effects model, the random effects variance is accounted for by including an extra error component.

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \mu_i + \varepsilon_i \quad (2.8)$$

The coefficients in this model are calculated via weighted least squares algorithms, with the weights defined by the inverse of the estimated variances:  $w_i = 1/v_i$  with  $v_i = \text{var}(\mu_i + \varepsilon_i)$

The meta-analysis models for combining effect sizes discussed in this section can be classified as shown in Table 2.1.

**Table 2.1:** A two-dimensional classification of meta-analysis models based on the characteristics of the distribution of effect sizes.

	No systematic variation	Systematic variation
No random variation beyond sampling error	Fixed effects model	Fixed effects regression model
Random variation beyond sampling error	Random effects model	Mixed effects regression model

Both the fixed effects regression model and the mixed effects regression model are estimated by means of meta-regression analysis<sup>4</sup>. Because of the largely non-experimental set-up in economic research, the sources of systematic heterogeneity among studies are greater in number than they are, for example, in the experimental sciences. Therefore, meta-regression

<sup>4</sup> Note that the fixed effects model and the random effects model can be considered as meta-regression models in which the effect size is regressed on a constant only.

analysis is a relatively important technique in the economic meta-analysis literature. In a meta-regression analysis, the effect size is regressed on a number of moderator variables that can be study characteristics or other conditioning variables. Moderator variables are often based on the following groups of study characteristics (from Florax and de Groot, 2002).

- theory and methodology (modeling framework, operationalizations)
- temporal dynamics (time trend, time period)
- spatio-temporal dynamics (geographical information and time span data)
- estimation details (functional form, estimation method)
- type of data (time-series vs. cross-section, frequency, aggregation level)
- publication details (published vs. unpublished, publication type, quality of publication outlet)
- quality of the study (theory, methods, data)

Typically, most of the moderator variables are dummy variables. By estimating the meta-regression model, the relationship between the variation in the effect size and the variation in the study characteristics can be identified. The aim of a meta-regression model is to combine effect sizes in order to obtain general mean estimates of the effect of interest, conditional on certain study characteristics. As such, it yields meaningful comparative results that allow for an investigation of the impact of study characteristics on the effect of interest. Furthermore, the results may facilitate cost-benefit assessment and economic forecasting through the possibility of (partial) transferability.

## **2.5 Potential Drawbacks and Methodological Issues**

Meta-analysis is not without its disadvantages and weaknesses. Three potential methodological problems can be distinguished, which are related to literature retrieval, heterogeneity among primary studies and multiple sampling (see Florax, 2002a).

Due to selection effects, the set of collected primary studies may insufficiently reflect the population of studies. *Selection bias* occurs when the process of literature retrieval is such that the probability of including a primary study in the meta-analysis is affected by certain conditioning factors. Selection bias can be caused by a meta-analyst's focus on a specific theoretical or modeling approach or by restrictive sampling by the meta-analyst over time,

within a country or within a language zone. Another important cause of selection bias is *publication bias*. Publication bias occurs when the probability of publication of a study is affected by certain conditioning factors so that published studies may not be representative of the population of studies. It has long been accepted that research with statistically significant results is potentially more likely to be submitted, published or published more rapidly than work with null or non-significant results (Sutton et al., 2000). Both editorial selection processes and self-censoring by the author may be the drivers behind this. Other factors that were found to affect publication probability are sample size, preferences for certain study results, presence or absence of randomization, exploratory versus confirmatory studies calendar time, the nature of a journal, source of funding and study language (see Card and Krueger, 1995; Dickersin et al., 1987; Begg and Berlin, 1988). Publication bias can be detected by visual methods or by statistical tests. The basic idea underlying every detection method is that a systematic relationship between the effect size and its standard error (or sample size) indicates the presence of publication bias. For an overview of methods we refer to Sutton et al. (2000, pp.112-119) or Stanley and Jarrell (2005). If publication bias is detected, this should be taken into account in the study retrieval phase or during the actual analysis. First, efforts can be focused on the retrieval of unpublished studies. In this case, the publication bias is not remedied but the degree to which it leads to selection bias in the dataset may be reduced. Furthermore, Sutton et al. (2000, pp. 119-125) mention several methods that correct for publication bias by adjusting the results of a meta-analysis. However, none of these methods is carried out routinely, and it remains to be seen how useful and appropriate they are in practice.

Due to the largely non-experimental character of economic science, economic studies exhibit a fair amount of heterogeneity. Primary studies may exhibit heterogeneity with respect to many factors such as locational setting, time period and model specification. If these factors are observed and measurable, then the heterogeneity can be addressed by adequate modeling in a meta-regression model. If part of the heterogeneity stems from unobserved factors, a mixed effects model can be used to correct for it. The methods to deal with these cases of heterogeneity have been extensively discussed in Section 3. Heterogeneity becomes a problem if the difference between studies is such that it renders the effect sizes incomparable to each other.<sup>5</sup> Incidentally, it is not always obvious whether effect size estimates are comparable or can be made comparable. For example, the direct comparability of estimates

---

<sup>5</sup> Sometimes, effect sizes can be made comparable by simple transformation or rescaling, or by more complex re-computations which may require information on the underlying data of the effect size estimates.

based on point elasticities and those based on arc elasticities may be questioned, as these elasticities do not measure the exact same effect (see Koetse et al., 2005). A further issue entails the question whether the effect sizes can be made comparable by the inclusion of a dummy variable in the meta-regression analysis or if point and arc elasticities are too fundamentally different for such an approach. The way of handling these issues may depend on the precise research questions and the assumptions of the analyst.

A third methodological problem in meta-analysis is related to multiple sampling, i.e. the collection of multiple observations per study. Since observations from the same study have study characteristics in common, they are statistically correlated. This causes within-study dependence in the data<sup>6</sup>. Part of the dependence is observed and can be addressed by proper modeling in a meta-regression framework. However, part of the dependence is unobserved and thus picked up by the disturbance term. Hence, the assumption of non-autocorrelation is violated and the covariance matrix of the disturbance term is non-spherical. The problem is comparable to the problem of serial correlation in time-series econometrics and has similar consequences; the OLS estimator is no longer efficient, although it is still unbiased. Therefore, statistical inference based on the routinely computed variance and standard errors will be misleading.

Two ways of dealing with the problem of within-study dependence can be identified. The first is to use conventional estimation techniques such as OLS and subsequently adjust the results to correct for the dependence between observations. The second is to use estimation techniques that account for dependence within the data set. While such estimation techniques are not specifically designed to deal with dependence in meta-analysis, and therefore not widely applied in meta-analytical studies, they may still provide good alternatives to the estimation techniques that are usually applied in meta-analysis.

## 2.6 Meta-analysis in Economic Research

The assets of meta-analysis as a research method to study issues in transport economics are discussed in Section 1.3; by pooling study results together, more reliable and precise estimates of a variable of interest, such as a valuation index or a price elasticity, can be obtained which are not constrained by the characteristics of a single empirical study. Furthermore, meta-analysis enables the analyst to estimate the impact value and significance

---

<sup>6</sup> This also holds, to a lesser degree, for observations that come from different studies that have very similar characteristics.



of conditioning factors on the variable of interest. These results could for example be used to estimate the “expected” value of the variable of interest in a specific setting. As such, meta-analysis may facilitate value transfer in cost-benefit analysis or calibration of CGE models.<sup>7</sup> In this section, we focus on certain characteristics of economic research and their effect on the properties of the dataset and the estimation procedure in meta-analysis.

In the economic meta-analytical literature, the effect size typically has the form of a (point) elasticity or a nominal value, such as consumer surplus or willingness to pay. The advantage of using point elasticities is that the effect size does not depend on the unit of measurement, and its distribution is known to be asymptotically normal (Florax, 2002a). Due to the largely non-experimental set-up in economic research, the sources of heterogeneity among studies are greater in number than they are in the experimental sciences, which causes systematic variation in the effect sizes. Therefore, meta-regression analysis is usually the analysis technique of choice, often in combination with descriptive statistics. The choice between a fixed effects regression model and a mixed effects model can be based on a homogeneity test on the effect size variation that remains after correcting for the systematic variation.

As in the experimental sciences, publication bias is a potential problem in meta-analyses in economics. Studies on publication bias, detection methods and remedies are generally based on a single sampling set-up and generally assume a single true underlying effect size (see for example Aschenfelter, 1999; Florax, 2002b; Stanley, 2005). How publication bias might work in a set-up where the collection of effect sizes is based on multiple sampling techniques and where there is systematic variation in effect sizes, is still relatively under-researched<sup>8</sup>. Hence, the validity of the theory, detection methods and remedies discussed in these studies under a multiple sampling scheme and under systematic variation in the effect size, which are common characteristics of meta-analysis in economics, is uncertain and needs further research. We leave this issue for further research.

---

<sup>7</sup> Value transfer refers to the use of the monetary value of a good obtained in a given context (often called study site), to estimate the value of a similar good under a different context (policy site) (Desvouges et al. 1992). Computable general equilibrium (CGE) modelling generally aims to build a model with a relatively transparent structure in order to clarify the mechanism with which policy measures or exogenous shocks affect the economy within a multisector framework (Thissen 1999). CGE modelling involves a calibration phase in which the values of model parameters are determined.

<sup>8</sup> In a multiple sampling set-up, one might expect within-study dependence to be present in both the effect size estimates and the standard errors. In such a case, publication bias detection methods, which focus on a systematic relationship between the effect size and its standard error (or sample size), may indicate the presence of publication bias, while the systematic relationship might actually be caused by within-study dependence.

The problem of lacking independence has not been addressed sufficiently in economic meta-analyses (Florax, 2002a). Few studies refer to the potentially disturbing influence of correlated effect sizes, although the occurrence of dependence is much more likely in economics - as compared to, for instance, medicine - because of multiple sampling of estimates per study. The problem is aggravated due to the relatively large amount of heterogeneity between economic studies. In fact, heterogeneity in effect sizes and correlation between effect sizes from the same study are closely linked and difficult to distinguish from each other. Under multiple sampling conditions, any variation in effect sizes is partly caused by within study correlation. In such a situation, based on the results of a  $Q$ -test, an estimation model might be used that accounts for heterogeneity (i.e. a random effects model or a mixed effects model) but which ignores the multiple sampling structure of the dataset and thus does not correct for within-study dependency. Despite the fact that this may lead to inefficient estimation results, statistical testing for dependence is virtually non-existent in economic meta-analysis.

## 2.7 Conclusion

In this chapter, we describe the background and characteristics of meta-analysis as a research method. We start by describing the history and background of meta-analysis by means of a brief literature overview. Subsequently, we discuss the methodology and strength of meta-analysis as a research method. We show that the meta-analytical method has several advantages on traditional qualitative literature review. Next, we discuss a number of estimation models used in meta-analysis, i.e. the fixed effects model, the random effects model, the fixed effects regression model and the mixed effects regression model. The choice of estimation model depends on the presence of systematic or non-systematic variation in the effect sizes. In the case of systematic variation, which is common in the economic literature, a multivariate model should be used. In the next section, we point to a number of drawbacks and methodological issues of meta-analysis. We show that selection and publication bias, heterogeneity and multiple sampling are potential sources of errors in the estimation results. Finally, we discuss a number of characteristics of economic research that have consequences for the properties of the dataset and the estimation procedure in meta-analysis. Publication bias is difficult to detect and correct for in the typical set-up of economic meta-analysis, which is characterized by multiple sampling and large heterogeneity between studies. Due to

the large heterogeneity, multivariate techniques such as meta-regression analysis are commonly used. While both heterogeneity and within-study correlation can be detected by tests and corrected for by estimation techniques, it is difficult to distinguish between these two properties of the dataset.

Based on the strengths, methodology and empirical applications of meta-analysis that we discuss in this chapter we conclude that there is a lot of potential for the application of meta-analysis in the field of transport economics. However, we also point to the fact that the estimation difficulties associated with multiple sampling are particularly relevant in meta-analyses in economics and therefore deserve attention before conducting applied meta-analyses in the field of transport economics. Hence, in the next chapter we investigate the issue of within-study dependence in more detail in the next chapter by means of a Monte Carlo study. In this simulation study, we compare the performance of various estimation techniques in the presence of within-study dependence in the data.

## Chapter 3

# Accounting for Multiple Sampling in Meta-Analysis. A Monte Carlo Study of Estimator Performance

### 3.1 Introduction

In the previous chapters, we discussed two important aims of the application of meta-analysis, i.e., (i) the combining of effect sizes by means of a fixed effects or random effects model, depending on the exact distributional assumptions regarding the true effect size and (ii) the identification of the impact of study characteristics on the effect size by means of regression analysis<sup>9</sup>. In a meta-regression analysis, the effect size is regressed on a selection of study characteristics and other explanatory variables. If the assumptions of the classical linear regression model hold, the estimation of the regression model can be estimated with OLS<sup>10</sup>. However, the assumption of non-autocorrelation is likely to be violated in a meta-analytical setup. The reason for this is that studies often provide more than one observation. Observations from the same study necessarily share some of the same underlying data. This results in within-study dependence, i.e., dependence between observations from the same study.

In this chapter, we aim to investigate and compare the performance of four different estimators in the case of within-study dependence. The study is structured as follows. In Section 3.2, we illustrate in detail the within-study dependence problem that pervades meta-analysis and introduce a number of models, designed to deal with similar dependence patterns. Section 3.3 describes the experimental design of the simulation. In Section 3.4, we discuss the estimation models and the evaluation criteria that we use in our simulation. In contrast with Section 3.2, here we focus on aspects related to the actual estimation of the models. Section 3.5 discusses the results of the simulation. Section 3.6 concludes.

---

<sup>9</sup> The fixed effects model assumes only sampling error. The random effects model assumes also non-systematic variation in the true effect size across studies. The meta-regression model is used in the case of systematic variation in the true effect size. It can be estimated under fixed or under random effects assumptions (the latter is called a mixed effects model). For an overview of the different models and how they are related see Figure 2.1.

<sup>10</sup> The assumptions can be found in most econometric textbooks. See for example Greene (2003).

### 3.2 Dealing with Within-Study Dependence in Meta-Analysis

As we described in Chapter 2 of this thesis, meta-regression analysis is a multivariate meta-analytical technique that is used to investigate systematic variation in effect size estimates found in the literature. A simple meta-regression model can be written as follows:

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + \mu_i \quad (3.1)$$

where  $y_i$  denotes the observed effect size,  $\mathbf{x}_i$  is a vector with data on moderator variables and  $\mu_i$  denotes the disturbance term of observation  $i$ . Typically, most of the explanatory variables are dummy variables. Under the assumptions of the classical linear regression model we have:

$$\boldsymbol{\mu} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad (3.2)$$

and the OLS estimator is unbiased and efficient. However, the observations in a meta-regression database show variation in the values of the moderator variables that are used in the regression. A substantial number of moderator variables are variables that describe the data on the study level and as such, they have the same value for all observations that come from the same study. Hence, observations that originate from the same study are correlated. Part of the dependence is observed and can be addressed by adequate specification of the meta-model but some of the dependence is unobserved and thus picked up by the disturbance term. The assumption of non-autocorrelation is violated and we have that

$$\boldsymbol{\mu} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{\Omega}), \boldsymbol{\Omega} \neq \mathbf{I} \quad (3.3)$$

As a result, the OLS estimator is no longer efficient, although still unbiased. Statistical inference based on the routinely computed variance and standard errors will be misleading. If the structure of  $\boldsymbol{\Omega}$  is unknown, we can address this problem by basing our inference on Huber-White standard errors. Huber-White standard errors are adjusted for specified assumed-and-estimated correlations of error terms across observations as well as for heteroskedasticity. Estimation of Huber-White standard errors is robust under very general conditions.

However, in the case of meta-analysis we have information about  $\mathbf{\Omega}$ . We observe the number of studies and the number of observations per study. Within-study dependence in the disturbance terms results in a covariance matrix with blocks of non-zero elements along the diagonal. An example of such a covariance matrix of a small dataset of three studies with three, one and two observations, respectively, is shown in (3.4).

$$\sigma^2 \mathbf{\Omega} = \sigma^2 \begin{pmatrix} \omega_1 & \rho_{12} & \rho_{13} & 0 & 0 & 0 \\ \rho_{21} & \omega_2 & \rho_{23} & 0 & 0 & 0 \\ \rho_{31} & \rho_{32} & \omega_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \omega_5 & \rho_{56} \\ 0 & 0 & 0 & 0 & \rho_{65} & \omega_6 \end{pmatrix} \quad (3.4)$$

Here,  $\omega_i$  denotes the relative variance of observation  $i$ , while  $\rho_{ij}$  denotes the correlation between observations  $i$  and  $j$ . The number of blocks corresponds to the number of studies in the meta-analysis while the size of the blocks corresponds with the number of observations within each study.

The study-observation structure in meta-analysis suggests the use of a hierarchical modeling approach to the dependence problem.<sup>11</sup> Goldstein (1995) describes a two-level variance components estimator (MLVL) that is constructed from the key-notion that there is heterogeneity between studies as well as within studies. Henceforth the model contains two disturbance components: one on the observation level and one on the study level:

$$y_{is} = \mathbf{x}'_{is} \boldsymbol{\beta} + \varepsilon_{is} + \mu_s \quad (3.5)$$

where  $i = 1 \dots N$  indexes the observations and  $s = 1 \dots S$  indexes the studies. This results in a nested structure in the covariance matrix of the disturbance. The covariance matrix has the same block-diagonal structure as in (3.4). In the case of our example of a dataset with six observations the covariance matrix would be:

---

<sup>11</sup> For an introduction to hierarchical modelling and the relation with meta-analysis see Raudenbush (2002) or Goldstein (1995).

$$\mathbf{V} = \sigma^2 \mathbf{\Omega} = \begin{pmatrix} \sigma_\varepsilon^2 + \sigma_\mu^2 & \sigma_\mu^2 & \sigma_\mu^2 & 0 & 0 & 0 \\ \sigma_\mu^2 & \sigma_\varepsilon^2 + \sigma_\mu^2 & \sigma_\mu^2 & 0 & 0 & 0 \\ \sigma_\mu^2 & \sigma_\mu^2 & \sigma_\varepsilon^2 + \sigma_\mu^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\varepsilon^2 + \sigma_\mu^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_\varepsilon^2 + \sigma_\mu^2 & \sigma_\mu^2 \\ 0 & 0 & 0 & 0 & \sigma_\mu^2 & \sigma_\varepsilon^2 + \sigma_\mu^2 \end{pmatrix} \quad (3.6)$$

From (3.4) we see that the dependence pattern is characterized by (i) multi-directionality, i.e., error terms can be correlated with more than one other error term and (ii) two-dimensionality, i.e., there is mutual dependence between pairs of correlated error terms (Anselin, 1988). As a result, the use of econometric techniques that deal with serial autocorrelation (such as found in the time-series literature) is not appropriate in the case of clustered autocorrelation.

Multi-directional and two-dimensional dependence between variables is a common issue in the spatial modeling literature (see for example Anselin and Tam Cho, 2002). The spatial error model (SEM) is a model that deals with such dependence in the error terms<sup>12</sup>. The model is written as

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i \quad (3.7)$$

$$\varepsilon_i = \lambda \mathbf{W} \varepsilon_i + \mu_i \quad (3.8)$$

The error term consists of a dependent part,  $\lambda \mathbf{W} \varepsilon$ , and an independent part,  $\mu$ . The disturbance terms are correlated with each other through a spatial weight matrix,  $\mathbf{W}$ , representing the topology of the spatial system and through  $\lambda$ , the scalar spatial autoregressive disturbance parameter<sup>13</sup>. From a meta-analytical perspective,  $\mathbf{W}$  represents the pattern of dependence, i.e., the number of studies and the number of observations of each study and  $\lambda$  represents the degree of within-study dependence. By rearranging terms the model can be expressed as

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + (\mathbf{I} - \lambda \mathbf{W})^{-1} \mu_i \quad (3.9)$$

<sup>12</sup> For an introduction to spatial econometrics see Anselin (1988).

<sup>13</sup> For a detailed discussion on the construction of spatial weight matrices and the application to meta-analytical data structures see Florax (2002).

### 3.3 Experimental Design

In the present simulation, we use a data generation process (DGP) to generate data sets that have the clustered correlation structure that characterizes data sets in meta-analysis, i.e., observed correlation between study characteristics and unobserved correlation between disturbances for observations that come from the same primary study. This means that we generate datasets with a clustered dependence structure in the explanatory variables and in the disturbance terms. We assume that within-study correlation pervades the observed and the unobserved part of the variation in observations to the same degree. Hence, in the datasets we generate, the degree of dependence in the explanatory variables is equal to the degree of dependence in the explanatory variables. Furthermore, we assume that the degree of within-study dependence, the variance in the disturbances and the variance in the explanatory variables are equal among primary studies.

#### 3.3.1 Generating Dependence

Our experimental design is based on the following model:

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i \quad (3.10)$$

$$\boldsymbol{\varepsilon} \sim N[0, \sigma_\varepsilon^2 \boldsymbol{\Omega}] \quad (3.11)$$

$$E[\mathbf{x}_k \mathbf{x}_k'] = \sigma_x^2 \boldsymbol{\Omega}, \forall k \quad (3.12)$$

where  $\boldsymbol{\Omega}$  is a block-diagonal matrix where the diagonal elements have value one and the off-diagonal elements within the blocks have value  $\rho$ , which specifies the degree of within-study correlation. An example for a meta-analysis dataset consisting of three studies with three, one and two observations, respectively, is shown in (3.13).

$$\boldsymbol{\Omega} = \begin{pmatrix} 1 & \rho & \rho & 0 & 0 & 0 \\ \rho & 1 & \rho & 0 & 0 & 0 \\ \rho & \rho & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \rho \\ 0 & 0 & 0 & 0 & \rho & 1 \end{pmatrix} \quad (3.13)$$



By varying  $\rho$ , we are able to generate datasets with different degrees of within-study correlation.

### 3.3.2 Variation in Design Factors

In order to generate a data set, certain factors of the simulation design, such as the number of studies in the data set, the number of observations per study and the degree of dependence, need to be determined. In order to be able to investigate the impact of design factors on estimator performance we generate datasets for different values of design factors (see Table 3.1) in order to induce systematic variation in the datasets.

**Table 3.1:** Variation in the values of the factors of simulation design

Design factor	Variation
Number of studies ( $N$ )	20, 30, 50
Average number of observations per study ( $N/s$ )	2, 3, 5
Variation in the number of observations per study	zero, low, high*
Lambda ( $\rho$ )	0, 0.1, ..., 0.9

\* In category "zero", all studies contain  $N/s$  observations. In category "low", 60 per cent of the studies contain  $N/s$  observations, 20 per cent contain 1 observation and 20 per cent contain  $2N/s - 1$  observations. In category "high", 60 per cent of the studies contain 1 observation, 20 per cent contain  $N/s$  observations and 20 per cent contain  $4N/s - 3$  observations.

Note that the first three design factors determine the size and structure of the covariance matrix and thus the pattern of dependence while the fourth design factor, lambda, determines the degree of dependence. We generate data sets for each of the 270 possible combinations of design factor values. For each combination of design factors, we generate 1000 data sets.

### 3.3.3 The Data Generating Process in Steps

In order to investigate the performance of the estimators we perform a simulation in which the true value of the effect size, as well as the impact of the moderator values are known. Table 3.2 lists the variables and parameters that are involved and shows their purpose in the data-generating phase and in the estimation phase.

In order to generate a data set we follow a number of steps. First, we set the true values of four betas at 0 and the true values of four other betas at 1. Next, we determine the values of the design factors. The third step is to randomly generate four vectors of moderator

variables that are uniformly distributed on the interval  $[0,1]$  with  $E[\mathbf{x}_k \mathbf{x}_k'] = \sigma_x^2 \mathbf{\Omega}, \forall k$  and four vectors of dichotomous variables with equal probability being 0 or 1 with  $E[\mathbf{x}_k \mathbf{x}_k'] = \sigma_x^2 \mathbf{\Omega}, \forall k$ . We choose  $\sigma_x^2$  equal to 1. With  $\mathbf{X}$  and  $\beta$  known we can compute the vector of true effect sizes as  $\mathbf{y}^{true} = \mathbf{X}\beta$ . The last step is the generation of the vector of dependent error terms  $\boldsymbol{\varepsilon} \sim N[\mathbf{0}, \sigma_\varepsilon^2 \mathbf{\Omega}]$ . The observed effect sizes  $\mathbf{y}^{obs}$  are now calculated as  $\mathbf{y}^{obs} = \mathbf{X}\beta + \boldsymbol{\varepsilon}$ .

**Table 3.2:** The implementation of various variables and parameters in the data generation and estimation phase

Symbol	Description	Data generation phase	Estimation phase
$\beta$	Coefficients	fixed	estimated
$\rho$	Degree of dependency	fixed	estimated
$\mathbf{X}$	Moderator variables	generated randomly but with clustered dependency based on $\rho$	observed
$\boldsymbol{\varepsilon}$	Disturbances	generated randomly but with clustered dependency based on $\rho$	unobserved
$\mathbf{y}^{obs}$	Effect size	computed as a function of $\mathbf{X}$ , $\beta$ and $\boldsymbol{\varepsilon}$	observed

### 3.4 Estimation Methods and Evaluation Criteria

The aim of this chapter is to compare the performance of the four estimation methods introduced in Section 3.2. In this section, we describe the actual estimation process for each of these models. Furthermore, we discuss the evaluation criteria that we use to compare the performance of these models.

#### 3.4.1 Estimation Methods

The estimation methods introduced in section 3.2 are (i) OLS, (ii) OLS with Huber–White corrected standard errors, (iii) the spatial error model and (iv) the 2-level variance components estimator. As discussed before, the estimated beta-coefficient under OLS is unbiased but inefficient because the autocorrelation violates the OLS assumptions. This problem can be addressed by using adjusted Huber-White standard errors (Huber, 1967; White, 1980) for inference. Huber-White standard errors can be derived after estimating the variance matrix of estimated coefficients as follows:

$$\text{Var}[\mathbf{b}] = \frac{1}{n} \left( \frac{1}{n} \mathbf{X}'\mathbf{X} \right)^{-1} \hat{Q}_* \left( \frac{1}{n} \mathbf{X}'\mathbf{X} \right)^{-1} \quad (3.14)$$

with

$$\hat{Q}_* = \left( \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n e_{is} e_{jt} x_{is} x'_{jt} D(s, t) \right) \quad (3.15)$$

where,  $D(s, t)$  is an indicator function that has value 1 if the study indexes  $s$  and  $t$  are equal and value 0 otherwise.

The spatial error model accounts for the dependence structure of the data by the use of a spatial weight matrix  $W$  and an auto-regression parameter,  $\lambda$ . The likelihood function is based on the assumptions of the spatial error model. This means that  $\mathbf{W}$  is known and the parameter  $\lambda$  is estimated together with the beta-coefficients  $\beta$  and the standard errors  $e$ . The log-likelihood function to be estimated is:

$$\begin{aligned} L = & \sum_i \ln(1 - \lambda \omega_i) - \frac{N}{2} \ln 2\pi - \frac{N}{2} \ln \sigma^2 \\ & - \frac{1}{2} \sigma^2 \left[ (\mathbf{y} - \mathbf{X}\beta)' (\mathbf{I} - \lambda \mathbf{W})' (\mathbf{I} - \lambda \mathbf{W}) (\mathbf{y} - \mathbf{X}\beta) \right] \end{aligned} \quad (3.16)$$

Estimates of  $\beta$  and  $\sigma^2$  can be expressed in  $\lambda$  as follows:

$$\mathbf{b} = \left[ (\mathbf{X} - \lambda \mathbf{W}\mathbf{X})' (\mathbf{X} - \lambda \mathbf{W}\mathbf{X}) \right]^{-1} (\mathbf{X} - \lambda \mathbf{W}\mathbf{X})' (\mathbf{y} - \lambda \mathbf{W}\mathbf{y}) \quad (3.17)$$

$$\mathbf{s}^2 = \frac{(\mathbf{y} - \mathbf{X}\mathbf{b})' (\mathbf{I} - \lambda \mathbf{W})' (\mathbf{I} - \lambda \mathbf{W}) (\mathbf{y} - \mathbf{X}\mathbf{b})}{N} \quad (3.18)$$

By substituting  $\mathbf{b}$  and  $\mathbf{s}^2$  into (3.16) the problem simplifies to the maximization of the concentrated likelihood function in  $\lambda$ :

$$L_c = -\frac{N}{2} \cdot \ln \left( \frac{\mathbf{e}'\mathbf{e}}{N} \right) + \sum_i \ln(1 - \lambda \omega_i) \quad (3.19)$$

with

$$\mathbf{e} = (\mathbf{y} - \mathbf{X}\lambda\mathbf{W}\mathbf{y})(\mathbf{y} - \lambda\mathbf{W}\mathbf{y})' (\mathbf{X} - \lambda\mathbf{W}\mathbf{X}) \times \left[ (\mathbf{X} - \lambda\mathbf{W}\mathbf{X})' (\mathbf{X} - \lambda\mathbf{W}\mathbf{X}) \right]^{-1} (\mathbf{X} - \lambda\mathbf{W}\mathbf{X})' (\mathbf{y} - \lambda\mathbf{W}\mathbf{y}) \quad (3.20)$$

The concentrated likelihood function is maximized using a bisection search method (see Anselin and Hudak, 1992).

The fourth estimation model we use is the 2-level variance components estimator, shown in (3.21).

$$y_{is} = \mathbf{x}'_{is}\boldsymbol{\beta} + \varepsilon_{is} + \mu_s \quad (3.21)$$

The model is estimated by recursive estimation of  $\boldsymbol{\beta}$  and the covariance matrix  $\mathbf{V}$ , given in (3.6). For the first iteration  $\sigma_\varepsilon^2$  and  $\sigma_\mu^2$  are both set at one. First  $\boldsymbol{\beta}$  is estimated by regressing  $\mathbf{y}$  on  $\mathbf{X}$ , using GLS with  $\mathbf{V}$  as a weight matrix:  $\mathbf{b} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ . Then  $\mathbf{V}$  is estimated by regressing the vectorization of the covariance matrix of the residuals  $\mathbf{ee}'$  on a set  $\mathbf{Z}$  of two dummy variables that have the value one if the corresponding cell in  $\mathbf{V}$  contains  $\sigma_\varepsilon^2$  or  $\sigma_\mu^2$ , respectively. For this estimation, GLS-estimation is used, with  $\mathbf{V}^* = \mathbf{V} \otimes \mathbf{V}$  as a weight matrix:  $\sigma^2 = [\sigma_\varepsilon^2, \sigma_\mu^2] = (\mathbf{Z}'\mathbf{V}^{*-1}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{V}^{*-1}\text{vec}(\mathbf{ee}')$ . Next,  $\boldsymbol{\beta}$  is estimated again, based on the new values for  $\sigma_\varepsilon^2$  or  $\sigma_\mu^2$ . In this fashion,  $\boldsymbol{\beta}$  and  $\mathbf{V}$  are recursively estimated until the estimated values for  $\boldsymbol{\beta}$  converge. The degree of within-study correlation is given by  $\sigma_\mu^2/(\sigma_\mu^2 + \sigma_\varepsilon^2)$ .

### 3.4.2 Evaluation Criteria

In order to compare the performance of the estimators we use the following three evaluation criteria: (i) bias (ii) variance and (iii) significance. Coefficient bias is calculated as the expected difference between the estimated betas and the true betas:

$$\text{Bias}(b_k) = E(b_k - \beta_k) \approx \frac{1}{n} \sum_{i=1}^n (b_{ik} - \beta_{ik}) \quad (3.22)$$

The efficiency is measured by the variance among the estimated coefficients. It is calculated as the mean squared difference between the estimated coefficients and the mean estimated coefficient:

$$\text{Var}(b_k) = E(b_k - \bar{b}_k)^2 \approx \frac{1}{n} \sum_{i=1}^n (b_{ik} - \bar{b}_{ik})^2 \quad (3.23)$$

We evaluate parameter significance by looking at the probabilities of type I and type II errors. A type I error occurs when a true null-hypothesis is incorrectly rejected; a type II error occurs when a false null-hypothesis is incorrectly accepted. The probability of a type I error is determined by the significance level, which is under the control of the analyst. In this study, we use a significance level of five percent. The probability of a type I error is equal to:

$$P(\text{type I}) = \frac{1}{n} C \left\{ i \mid |t_i| > t_{\alpha/2, n-k} \right\} \quad (3.24)$$

where  $C$  is an operator that denotes the number of elements in a set and  $\alpha$  specifies the two-sided significance level. We estimate type I error probabilities for the true betas that are equal to zero. For a given significance level, one would like the probability of type II errors to be as low as possible. The probability of a type II error is calculated as:

$$P(\text{type II}) = 1 - \frac{1}{n} C \left\{ i \mid |t_i| > t_{\alpha/2, n-k} \right\} \quad (3.25)$$

We estimate type II error probabilities for the true betas that are equal to one.

### 3.5 Simulation Results

In the data generation phase we have generated 1,000 datasets for each of the 270 combinations of design factors. We carried out regression estimations for all 270,000 datasets. Based on the regression results we calculated the following performance indicators: average estimation bias, the average estimation variance and the probabilities of type I and type II errors for each combination of design factors. Using different estimation models for the

regression analysis enables us to compare their performances. Furthermore, the results enable us to investigate how the performances are affected by changes in simulation design factors.

### 3.5.1 Estimator Performance

Figure 3.1 shows the average estimation bias of OLS, SEM and MLVL<sup>14</sup> for different levels of within-study correlation<sup>15</sup>. The estimation bias is zero for all the estimators and is unaffected by the degree of dependence.

Figure 3.2 shows the average estimation variance of OLS, SEM and 2LV for various levels of dependence. The results are quite different for all three estimators. In the absence of within-study correlation, the estimation variances of OLS and MLVL are similar (0.014) and that of SEM is slightly higher (0.016). When the degree of correlation increases, the variance of SEM increases seemingly linearly at first and non-linearly if the degree of correlation exceeds 0.7. The variance of OLS increases non-linearly over the whole range and exceeds that of SEM for high degrees of correlation. The increase in the variance of MLVL is very small compared to the other estimators. Clearly, the multilevel estimator outperforms the other estimators, except in the absence of within-study dependence.

Figure 3.3 shows the average probability of a type I error for OLS, HW, SEM and MLVL. The pattern is quite similar to that of the estimator variance. The values for all estimators except for MLVL increase in the dependence level, which indicates underestimation of the standard error. This results in too short confidence intervals when using any of these estimators. The value for OLS increases non-linearly in the dependence level. The value for SEM increases linearly at first and non-linearly for very high levels of dependence. The value for MLVL increases in the dependence level but, as in Figure 3.2, this increase is very small compared to that of the other two estimators. Inference based on Huber-White standard errors leads to an even higher probability of type I errors than OLS. Since the coefficient values are identical, this must be caused by lower standard errors for HW. In this case, one would expect a lower average type II error probability for HW.

Figure 3.4 shows that this is indeed the case; HW performs better than OLS. It also shows that MLVL again outperforms all the other estimators. The performance of SEM with

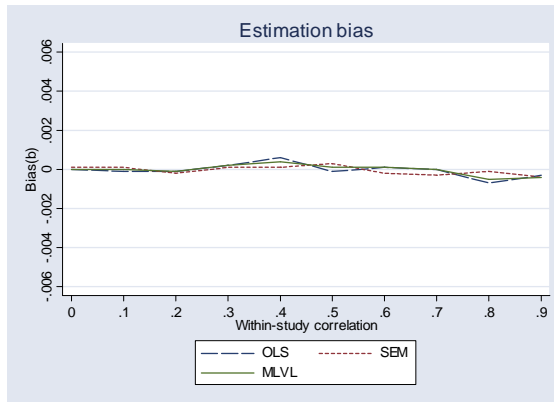
<sup>14</sup> Huber-White correction does not affect the estimation bias or the estimator variance. Therefore, with respect to these performance indicators, the results of Huber-White are equal to the results of OLS.

<sup>15</sup> For each degree of within-study dependence, the estimation bias is averaged over all true beta's and all generated datasets (i.e., all combinations of number of studies, number of observations per study and variation in the number of observations per study). The same holds for the estimation variance and the type I and type II error probability in Figure 3.2-3.4.

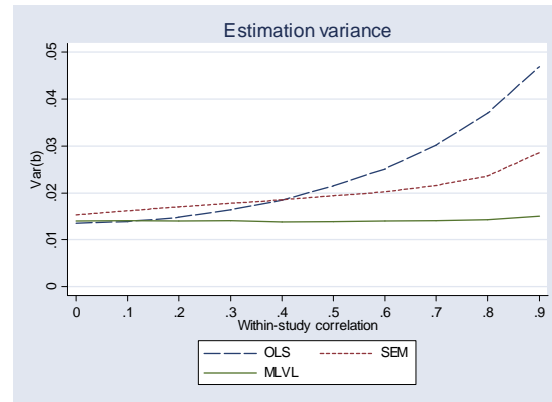
respect to type II error probability is worst for within-study correlation below 0.7 and second best for higher dependence levels.

### 3.5.2 The Effect of Design Factors on Estimator Performance

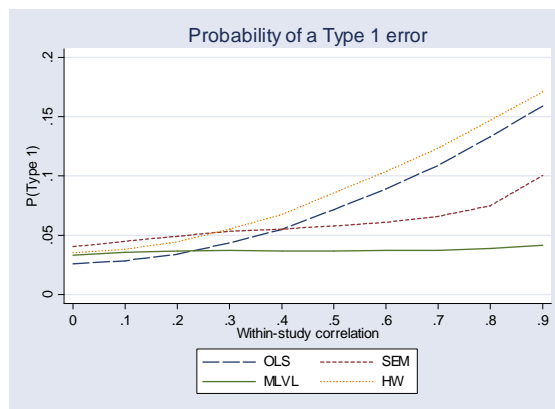
While it is useful to compare the performance of different estimators in the presence of increasing within-study dependence, it is equally interesting to investigate if the relationship



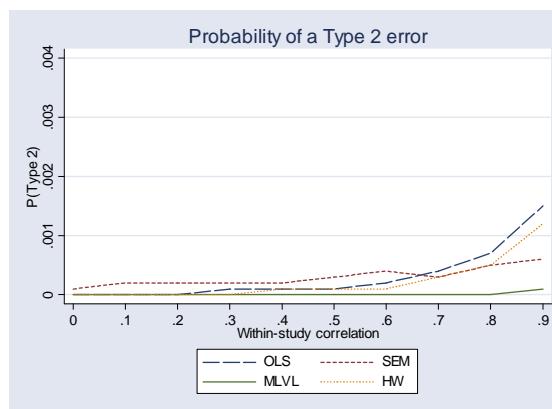
**Figure 3.1:** The estimation bias for different levels of within-study correlation



**Figure 3.2:** The estimation variance for different levels of within-study correlation



**Figure 3.3:** The probability of a type I error for different levels of within-study correlation



**Figure 3.4:** Probability of a type II error for different levels of within-study correlation

between within-study dependence and performance is affected by simulation design factors. Identification of such factors enables one to make a more informed choice of estimation method, based on knowledge about the structure of the dataset in terms of the number of

studies that have been used, the number of observations per study and so on. In the following section, we discuss the impact of simulation design factors on the estimation variance.

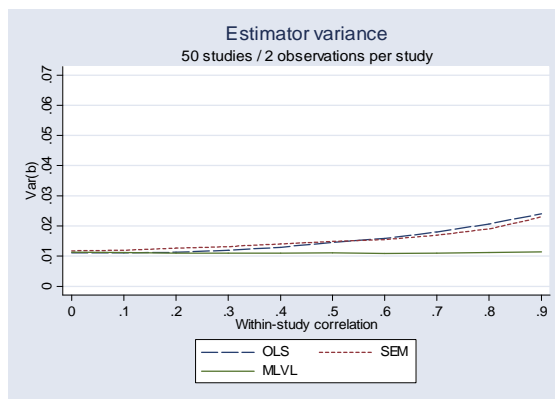
Figure 3.5 shows the estimator variance of OLS, SEM and MLVL based on the sub-sample of datasets that have 50 studies and an average of 2 observations per study, while Figure 3.6 shows the estimator variances of the estimators based on the sub-sample of datasets that have 20 studies and an average of 5 observations per study. Comparing these two figures allows us to investigate the impact of the average number of observations per study on the relationship between dependence level and estimation variance.

As expected, in the case of zero dependence the variances for both graphs are similar. If we increase the dependence, we see that the variance of MLVL increases only slightly in both graphs. The variances for OLS and SEM increase in both graphs, but not in the same fashion. In the left-hand graph, the estimator variances of OLS and SEM increase to about twice that of MLVL. In the right graph the variance of SEM increases less while the variance of OLS increases to almost seven times that of MLVL. Apparently, the decrease in the performance of OLS, caused by within-study dependence, is even higher when the number of observations per study is high. This makes sense, since the total number of pair-wise correlations between observations from the same study increases in the average number of observations per study. Therefore, not only the degree of dependence (the correlation coefficient) but also the number of pair-wise correlated observations increases the estimation variance of OLS. On the other hand, MLVL and SEM seem to be able to account perfectly for this.

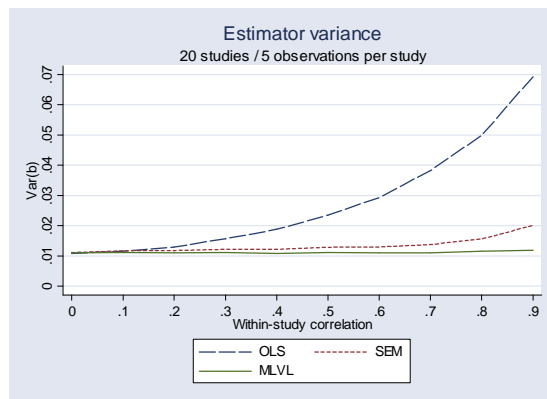
Figure 3.7 shows the estimator variance of OLS, SEM and MLVL based on the sub-sample of datasets that have "no" variation in the number of observations per study, while Figure 3.8 shows the estimator variances of the estimators based on the sub-sample of datasets that have "high" variation in the number of observations per study. Comparing the two figures allows us to investigate the effect of the variation in the number of observations per study on the relationship between dependence level and estimation variance. In the case of datasets where all studies have the same number of observations (Figure 3.7), the variance of OLS increases non-linearly in the degree of dependence up to a value of almost three times that of MLVL, while in the case of datasets with high variation in the number of observations per study (Figure 3.8), it increases up to a value of almost four times that of MLVL. The difference between the patterns in the two graphs seems to indicate that the negative effect of within-study dependence on the performance of OLS is exacerbated by an increase in the



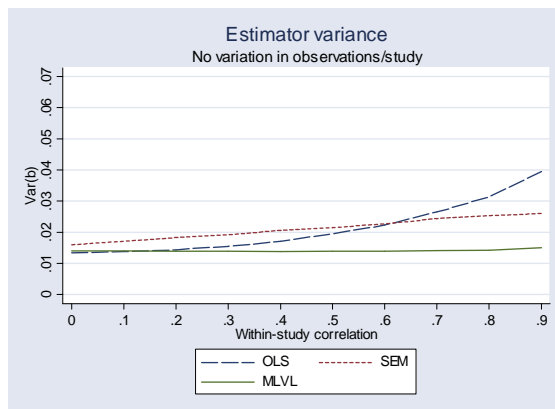
variation in the number of observations per study. In other words, if the distribution of observations over studies is highly unequal, the variance for OLS is more sensitive to an increase in variance. This makes sense for similar reasons as we discussed before with respect to the effect of the average number of observations per study. An increase in the variation in the number of observations per study is associated with an increase in the number of pair-wise correlations between observations. SEM seems to be able to account better for such variation than OLS and in fact performs even better in the case of high variation in the number of observations per study than in the case of datasets with no such variation. Finally, MLVL seems to have no problem at all handling either type of dataset.



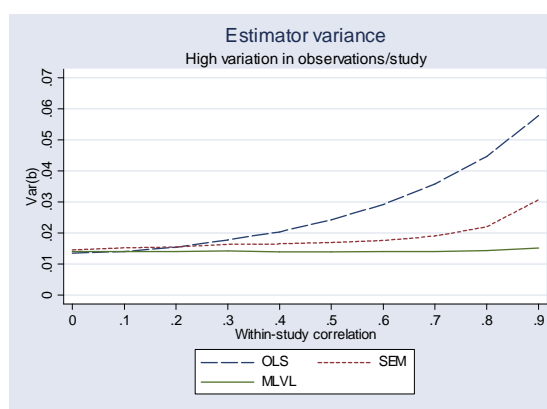
**Figure 3.5:** The estimator variance for datasets with an average of two observations per study



**Figure 3.6:** The estimator variance for datasets with an average of five observations per study



**Figure 3.7:** The estimator variance for datasets with no variation in the number of observations per study



**Figure 3.8:** The estimator variance for datasets with high variation in the number of observations per study

### 3.5.3 Response Surface

In Section 3.5.2, we investigated the effect of simulation design factors on the estimator variance of OLS, SEM and MLVL. In this section, we compare the estimation methods by investigating the effect of estimation method and design factors on estimator performance in a multivariate way. With respect to each performance indicator, i.e., estimator bias, estimator variance, type I error probability and type II error probability, we perform two regression analyses in which we treat the performance indicator as the dependent variable. In the first regression specification, we use the design factors as explanatory variables. Furthermore, using OLS as the reference category, we include dummy variables for SEM, MLVL and - if applicable - HW in order to check for level effects between estimation methods. In the second specification, we add interaction terms in order to investigate the effect of design factors on the difference in performance between estimation methods.

Column 1 and 2 in Table 3.3 show the results of two regression analyses with estimation bias as the dependent variable. In Column 1, we see that the degree of within-study correlation has a negative effect on the estimation bias. Even if the coefficient is so small that it would hardly affect the estimated coefficients of a meta-regression analysis in a way, this result is still a bit surprising. All other coefficients in Column 1 are insignificant, which indicates that the bias is not affected by the number of observations, the number of observations per study or the distribution of the observations over the studies. Furthermore, there is no significant difference between the estimation methods. These results are according to expectations. Column 2 shows that after adding interaction terms of estimation methods and design factors the effect of within-study correlation is rendered insignificant. All other coefficients also enter insignificant in this specification. These results confirm that with respect to estimator bias there is no difference in performance between OLS, SEM and MLVL and that all estimation methods are unbiased.

Column 3 and 4 show the results of two regression analyses with estimator variance as the dependent variable. The negative coefficients for SEM and MLVL in column 3 indicate that these estimation methods perform better than OLS. The difference in performance is largest for MLVL. Furthermore, we see that the estimation variance increases in the degree of within-study correlation and decreases in the number of observations, which is expected. The mean number of observations per study and the variation in the observations per study do not have a significant impact on estimator variance. Column 4 shows the results after including

interaction terms of the estimators and design factors. Again, the estimator variance increases in the degree of within-study correlation and decreases in the number of observations. Furthermore, we see that an increase in the mean number of observations per study increases the estimation variance. In addition, an increase in the number of observations per study produces higher levels of estimation variance. The added interaction terms allow us to investigate the estimator performance of SEM and MLVL in a detailed way.

**Table 3.3:** Response surface regressions on the estimator bias and estimator variance

Dependent variable	Estimator bias		Estimator variance	
	(1)	(2)	(3)	(4)
(Constant)	-0.132 (0.203)	-0.129 (0.327)	32.502 ** (1.292)	14.911 ** (1.770)
Degree of within-study correlation ( $\lambda$ )	-0.440 * (0.185)	-0.472 (0.321)	15.782 ** (1.177)	34.440 ** (1.736)
Number of observations	-0.001 (0.001)	-0.003 (0.002)	-0.136 ** (0.008)	-0.166 ** (0.011)
Observations per study	0.089 (0.058)	0.193 (0.101)	-0.348 (0.371)	2.824 ** (0.548)
Low variation in obs/study	0.244 (0.130)	0.181 (0.226)	0.037 (0.828)	1.316 (1.221)
High variation in obs/study	-0.002 (0.130)	-0.206 (0.226)	1.111 (0.828)	5.960 ** (1.221)
SEM	0.010 (0.130)	-0.036 (0.463)	-3.946 ** (0.828)	25.549 ** (2.503)
SEM*Correlation		0.095 (0.454)		-22.214 ** (2.455)
SEM*Nobs		0.004 (0.003)		0.025 (0.016)
SEM*Obs/study		-0.169 (0.143)		-5.552 ** (0.775)
SEM*Low var in obs/study		0.086 (0.320)		-2.503 (1.727)
SEM*High var in obs/study		0.332 (0.320)		-8.650 ** (1.727)
MLVL	0.017 (0.130)	0.055 (0.463)	-9.636 ** (0.828)	13.643 ** (2.503)
MLVL*Correlation		0.002 (0.454)		-33.762 ** (2.455)
MLVL*Nobs		0.003 (0.003)		0.068 ** (0.016)
MLVL*Obs/study		-0.141 (0.143)		-3.963 ** (0.775)
MLVL*Low var in obs/study		0.103 (0.320)		-1.332 (1.727)
MLVL*High var in obs/study		0.278 (0.320)		-5.898 ** (1.727)
N	810	810	810	810
R <sup>2</sup> -adjusted	0.007	-0.001	0.540	0.667

\* significant at the 5% level

\*\* significant at the 1% level

As expected, the degree of within-study correlation increases the difference in performance between OLS and the other two estimators. Apparently, SEM and MLVL handle the dependence better. The number of observations does not have a significant effect on the

difference in performance between OLS and SEM. The difference in performance between OLS and MLVL even decreases in the number of observations. However, the coefficient is so small that the superiority of MLVL over OLS is not contested. An increase in the sample size increases the performance of SEM and MLVL relative to OLS. This performance also increases if the variance in observations per study increases. The estimation results of this regression specification are consistent with Figures 3.1-3.4. For a detailed explanation of the results, we refer to that section. The results show that in the presence of within-study dependence, SEM and MLVL perform better than OLS with respect to estimator variance.

Column 1 and 2 in Table 3.4 show the results of the regressions with the probability of type I errors as the dependent variable. Column 1 shows that an increase in the degree of within-study dependence increases the type I error probability. The probability also increases in the number of observations per study and in the variation in the number of observations per study. This can be explained from an increase in the number of correlated observations. As we saw in the previous section, not only the degree of dependence but also the amount of pair-wise correlation between observations from the same study affects the estimator variance. Apparently, with respect to type I error probability there is a similar effect going on. Next, we see that an increase in the number of observations decreases the type I error probability. Comparing the performance of the different estimation methods, we see that SEM and MLVL both outperform OLS. The coefficients of the dummies for SEM and MLVL indicate that MLVL performs best. With respect to type I error probability, HW performs worse than OLS. As discussed before in Section 3.5.1, the reason is probably that, with equal coefficient estimates, Huber-White standard errors are lower than standard errors estimated with OLS. Column 2 shows the results after including the interaction terms. Focusing on the interaction terms, we see that the difference in performance between OLS and that of SEM and MLVL increases if (i) the degree of within-study correlation increases, (ii) the number of observations increases, (iii) the mean number of observations per study increases and (iv) the variation in the number of observations per study is high. HW performs better if the number of observations increases. Other design factors have no impact on the difference in performance between OLS and HW.

**Table 3.4:** Response surface regressions on the type I and type II error probabilities

Dependent variable	P (Type I error)		P (Type II error)	
	(1)	(2)	(3)	(4)
(Constant)	49.174 ** (0.597)	-53.682 ** (17.893)	50.190 ** (0.014)	52.288 ** (0.644)
Degree of within-study correlation ( $\lambda$ )	9.036 ** (0.315)	14.908 ** (0.370)	0.069 ** (0.007)	0.126 ** (0.013)
Number of observations	-2.607 ** (0.459)	0.019 (0.538)	-0.207 ** (0.011)	-0.260 ** (0.019)
Observations per study	0.610 ** (0.084)	1.723 ** (0.099)	0.003 (0.002)	0.003 (0.004)
Low variation in obs/study	0.280 (0.222)	0.600 * (0.260)	-0.002 (0.005)	0.006 (0.009)
High variation in obs/study	0.995 ** (0.222)	2.369 ** (0.260)	0.003 (0.005)	0.024 ** (0.009)
SEM	-1.462 ** (0.256)	126.554 ** (14.946)	-0.003 (0.006)	2.007 ** (0.538)
SEM*Correlation		-9.695 ** (0.523)		-0.076 ** (0.019)
SEM*Nobs		-5.759 ** (0.761)		-0.098 ** (0.027)
SEM*Obs/study		-2.370 ** (0.140)		0.001 (0.005)
SEM*Low var in obs/study		-0.647 (0.368)		-0.023 (0.013)
SEM*High var in obs/study		-2.842 ** (0.368)		-0.062 ** (0.013)
MLVL	-3.779 ** (0.256)	60.422 ** (14.946)	-0.029 ** (0.006)	-4.393 ** (0.538)
MLVL*Correlation		-14.297 ** (0.523)		-0.121 ** (0.019)
MLVL*Nobs		-2.536 ** (0.761)		0.223 ** (0.027)
MLVL*Obs/study		-1.904 ** (0.140)		-0.003 (0.005)
MLVL*Low var in obs/study		-0.606 (0.368)		-0.008 (0.013)
MLVL*High var in obs/study		-2.360 ** (0.368)		-0.025 (0.013)
HW	1.222 ** (0.256)	45.674 ** (14.946)	-0.008 (0.006)	-1.700 ** (0.538)
HW*Correlation		0.505 (0.523)		-0.028 (0.019)
HW*Nobs		-2.209 ** (0.761)		0.086 ** (0.027)
HW*Obs/study		-0.177 (0.140)		0.000 (0.005)
HW*Low var in obs/study		-0.025 (0.368)		-0.001 (0.013)
HW*High var in obs/study		-0.295 (0.368)		0.001 (0.013)
N	1080	1080	1080	1080
R <sup>2</sup> -adjusted	0.549	0.845	0.339	0.461

\* significant at the 5% level

\*\* significant at the 1% level

Column 3 and 4 show the results of two regression analyses with the probability of type II errors as the dependent variable. Column 3 shows that an increase in within-study dependence decreases the average performance of the estimators. Next, we see that for larger

datasets the probability of a type II error is lower than for small datasets. As before, MLVL performs significantly better than OLS. However, we find no such result for SEM and HW. The interaction terms in Column 4 show that the relative performance of both SEM and MLVL increases in the within-study dependence. Furthermore, we see that the type II error probability of SEM decreases if the variation in the number of observations per study is high and that SEM, MLVL and HW perform better if the number of observations increases. Other design factors have no effect on the relative performance of SEM, MLVL and HW.

### 3.6 Conclusion

The aim of this Monte-Carlo simulation study is to investigate and compare the performance of four different estimation methods with respect to the estimation of a meta-regression analysis in the case of within-study dependence. In section 3.2, we explain why meta-regression analyses are characterized by dependence among observations and we show that this dependence leads to a specific type of clustered autocorrelation. The problem of dependence in meta-analysis has similarities with the problem of spatial autocorrelation as described in the spatial econometrics literature. Furthermore, the datasets used in meta-analysis have a hierarchical structure in the sense that the sample of observations can be divided into groups, according to the primary study they originate from. Based on this information, we discuss three estimation models that we expect to perform better than OLS in the case of clustered autocorrelation because they account for the dependence structure in the data. In Section 3.3, we introduce the model that forms the basis of our experimental design and we described our data generation process in detail. Furthermore, we discuss a number of design factors that allow us to induce variation in the degree and the pattern of dependence in the generated datasets. Next, in Section 3.4 we describe the estimation process for each of the estimation models and the indicators that we employ as evaluation criteria upon which we base the performance comparison of the four estimation models.

In Section 3.5, we present the results of our simulation. We compared OLS, OLS with Huber-White corrected standard errors, the spatial autoregression model and a 2-level error components model (MLVL). The results show that all estimation models remain unbiased if we increase the degree of within-study dependence. However, with respect to estimator variance and type I and type II error probability MLVL outperforms the other models. The superiority of MLVL increases in the dependence level; only for very low levels of

dependence, the other models perform comparably. Furthermore, we find that not only the degree of dependence but also the pattern of dependence affects the performance of the estimators. An increase in the number of studies per observation and a more unequal distribution of observations over studies both lead to an increase in estimator variance for OLS. The variance of MLVL remains constant and is hardly affected by changes in the degree and the pattern of within-study dependence.

The results of this study support the use of multi-level estimation methods in meta-regression analyses. Unless the degree of within-study dependence is very low or altogether absent<sup>16</sup>, multi-level estimation is preferable over OLS. Based on the results of this chapter we suggest the following procedure when a meta-regression analysis is carried out. First, the presence and degree of within-study dependence are assessed by statistical testing. Goldstein (1995) suggests carrying out a likelihood ratio test by estimating the 'deviance' for the MLVL model and the model omitting the level 2 variance. The difference between the deviances is referred to tables of the chi-squared distribution with one degree of freedom. Furthermore, the result of MLVL estimation allows for calculation of a coefficient of within-study correlation. Based on the results of this test and the information about the observation-study structure of the dataset, the analyst can make an informed choice between the use of OLS or multi-level estimation. In the empirical Chapters 4, 5 and 6, where we apply meta-regression analysis techniques, we follow the procedure discussed in this paragraph.

---

<sup>16</sup> This could be the case, for example, if observation selection procedures are based on single sampling.

## **PART III**

# **APPLIED META-ANALYTICAL STUDIES IN TRANSPORTATION ECONOMICS**





## Chapter 4

# A Meta-Analysis of Price Elasticities of Demand for Aviation Travel<sup>17</sup>

### 4.1 Introduction

The aviation sector is in constant motion. The continuing growth in the number of passengers and aircraft movements necessitates a rise in investments in airport and aircraft capacity. However, even with these new investments, peak congestion and the environmental impact of aviation remain problematic. Air transport is apparently a field fraught with externalities. Another development in the aviation sector is the tendency to form alliances. Although the literature shows that these alliances may be beneficial to passengers, they still need consent from aviation authorities. In the deregulated aviation sector, aviation authorities therefore play a critical role in protecting the population from excessive noise and in safeguarding the consumer against “excessive” usage of market power.

As discussed in the introduction of this thesis, the price is a potentially powerful instrument of transport policy. For example, the government can put a price on the externality to reduce the negative effect, but if the passengers are not very sensitive to price changes, such a policy would have little effect; the airlines simply pass the charge on to the passengers. The government needs information on the price sensitivity of passengers in order to be able to estimate policy effects or to justify noise disturbance policy. This information is needed on different levels. For example, a kerosene tax can only be justified in the context of an international policy arrangement and requires different insights than a local noise charge.

However, the estimation of price elasticities in aviation is complicated by various problems concerning data availability on prices, numbers of passengers and so on. As an alternative, research synthesis of various empirical studies undertaken elsewhere or in the past could provide important insights on price sensitivity. Using existing research, one tries to find

---

<sup>17</sup> This chapter is based on Brons et al. (2002).

common factors explaining potential differences in e.g. estimates of price elasticity. This approach is followed in the present study.

In this study, we carry out a meta-analysis that investigates the variation in price elasticities of demand in the aviation sector. The purpose of this study is to test whether these price elasticity estimates encountered in the literature are statistically equal, and if not, to explain the variation in these elasticities by determining the impact of conditioning factors. The study is organized as follows. In section 4.2, we discuss the economic determinants of demand for air transport. Section 4.3 contains the empirical results and section 4.4 concludes.

## **4.2 Determinants of Demand for Passenger Air Transport**

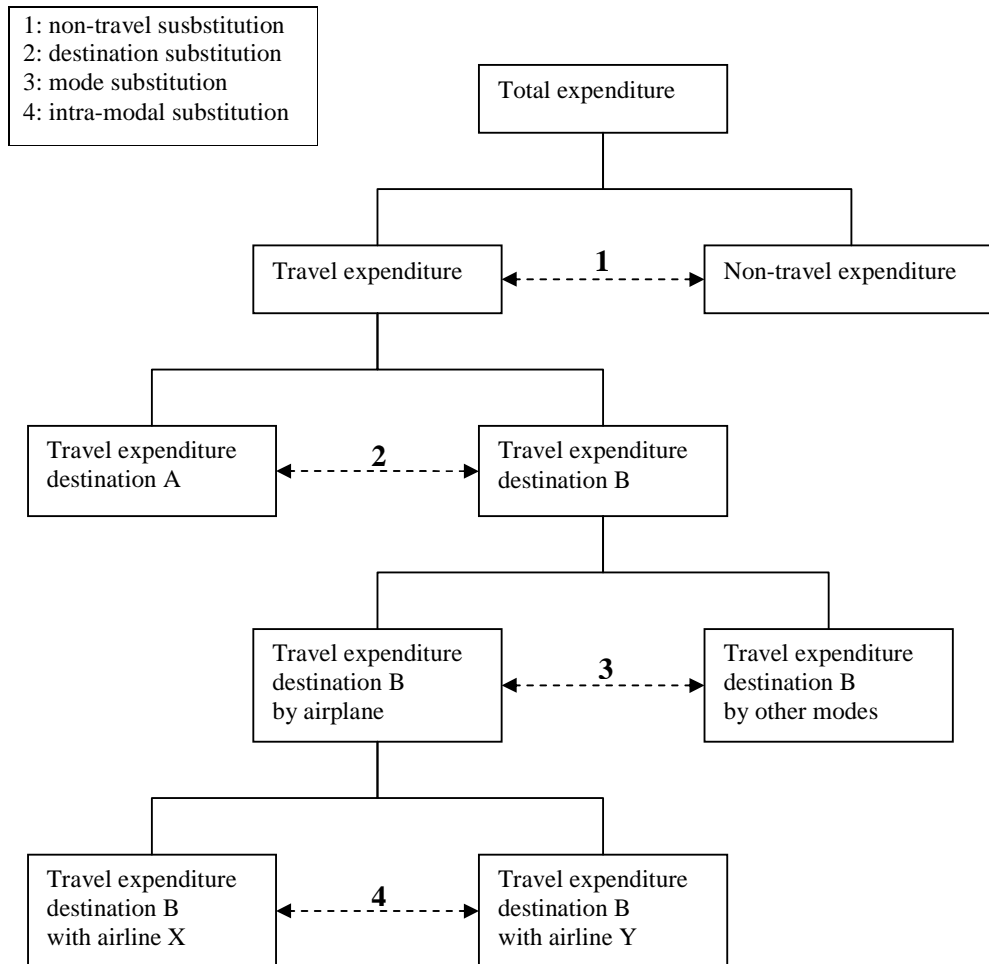
In this section, we discuss a number of economic, demographic and geographic determinants of the demand for passenger air transport. First, we discuss the relevance of substitutes for aviation transport. Subsequently, the determinants themselves, and the various ways in which they affect price elasticity levels are described. Finally, we discuss a number of study characteristic that may have an impact on the price elasticity. We conclude the section with a concise reflection.

### **4.2.1 Choice Contexts in Air Transport Demand**

The price elasticity of the demand for a consumer product or a production input factor is directly linked to the possibilities of substitution for that good. A relatively large number of substitutes results in a high price elasticity, whereas a lack of substitutes causes demand for a commodity to become inelastic. This also holds for the demand for passenger air transport. Most of the determinant factors of the price elasticity discussed in this study do not have a direct impact on the price elasticity, but affect the supply of substitution modes and thus have an indirect influence.

In the case of aviation transport, multiple levels of substitution can be distinguished, as shown by Figure 4.1. First, different carriers compete with each other on the same route, providing a case of intra-modal substitution. In the case of homogeneous transport services, the level of competition is higher, which implies higher price sensitivity of demand. On the other hand, when services of different quality are offered, demand will be more rigid. Next, on certain market segments, alternative transport modes that provide similar qualities can be

considered as substitution modes. Numerous factors, primarily of geographic, economic and demographic nature, determine the availability and the potential success of alternative modes as a substitute. It is obvious that geographic characteristics such as seas and mountain ranges and flight distance of a trip affect the possibilities of substitution.



**Figure 4.1:** Illustrative case of different levels of substitution in the aviation market

Third, destinations with similar characteristics are substitutes for each other. A useful way to look at this phenomenon is to adopt a hedonic point of view. The specific characteristics of a destination may be regarded as attributes, each of which contributes to some degree to the perceived overall utility of the destination. The rational consumer, subject to certain budget considerations, will choose the destination with the highest level of utility. If the relative ticket price for this destination increases, the consumer may start searching for destinations with a better price-to-utility ratio. Whether the consumer will succeed in finding a more favorable alternative depends on the availability of good substitutes for the original

destination. For instance, it is harder to substitute for an appointment with a business partner in New York than for a leisure trip to a generic Mediterranean seaside resort. Finally, due to budget constraints, non-transport goods can substitute transport goods if the utility level derived from buying the non-transport good are at least equal to the utility provided by the default consumption of the transport good; instead of booking a holiday trip, a consumer might choose to buy a new computer. Note that such trade-off effects are generally subject to diminishing rates of marginal return.

We conclude that transport prices play a role in various choice contexts. If the alternatives are homogeneous, we expect that price elasticities are ‘larger’ than when they are heterogeneous.

## 4.2.2 Economic Determinants of Demand and Price Sensitivity

Various economic, demographic and geographic factors may affect price sensitivity. The previous section showed that, from an economic point of view, price elasticities and substitution possibilities are strongly related. In fact, the impact of several determinant factors on the price elasticity is closely linked to the possibilities of substitution that are available.

High incomes are generally associated with a relatively high demand for air transport (see Mutti and Murai, 1977). In general, the income elasticity of demand for air transport is greater than unity, so that the income level and the *share* of aviation expenditures in the disposable income are positively correlated (see Crouch, 1991). This would suggest that, despite a decreasing marginal utility of income, the utility losses associated with a price increase are higher for this group of consumers, which would make them more price sensitive than consumers with low incomes.

As discussed in the previous section, possibilities for substitution are directly related to the distance of a flight. For long-distance flights, in particular intercontinental flights, less alternative transport modes are available. This suggests a negative relationship between distance and price sensitivity. On the other hand, long-distance flights are in general more expensive than short-distance flights, which means that a price increase would affect a larger share of the consumer’s budget.<sup>18</sup> This second aspect suggests a positive relationship between distance and price sensitivity. It is not clear which of the two effects will dominate.

---

<sup>18</sup> In today’s complicated aviation market, distance and fare price are not as closely related as one would expect. Landing rights, alliances, tariff classes and other market factors often dominate the cost structure of fares and distort such a close relationship (see for example Marsan and Kotoris, 1993 or Cooper and Maynard, 1971).

In addition to the travel distance, the geographic location may affect the price sensitivity. Cultural differences, availability of ground transport networks and difference in income can all have an impact on the price elasticity.

Leisure travelers, essentially consumers, aim to maximize the utility derived from the flight and the associated holiday experiences, subject to a given income or budget constraint. Characteristic determinants of leisure travel demand are travel costs, relative prices of other goods, income and socio-economic characteristics. Business travelers, who use a business flight as an input to final production, are in general interested in minimizing costs for a given level of output. Business travel demand is determined by such factors as travel costs, relative price of complementary production input factors and a firm's output level. Because of this, leisure and business travelers respond differently to changes in certain socioeconomic factors influencing the demand, and should therefore be modeled separately (Hooper, 1993).

There are several reasons why business travelers can be expected to be less price sensitive than leisure travelers. First, leisure travel is generally regarded as discretionary expenditure. Many goods and services compete with leisure travel for a share of the consumer's discretionary budget. Thus, for leisure travel there are more substitutes inside as well as outside of the transport sector than for business travel.

Furthermore, the generalized costs of travel include a value of time component. Since business travelers generally have a higher value of time than leisure travelers, the relative share of ticket prices in the total travel costs is low. This means that an increase in ticket price leads to a relatively small increase in the generalized costs of travel. The willingness to substitute monetary for temporal advantages is therefore lower for business class passengers. It is important, though, to observe the initial difference in ticket prices between business class and economy class trips. Since business class fares are generally much more expensive, the ticket price still forms a large part of the total travel costs for business class passengers and the monetary disadvantages of a price increase will exceed those of leisure travelers. The willingness to substitute monetary for time advantages may thus be higher than expected for business class passengers. The difference in price elasticity values between business class and economy class passengers therefore seems to depend on the difference between the ratio of the business class versus the economy class passengers' time valuation and the ratio of business class versus economy class initial fares.

---

Ample examples of not only decreasing marginal prices of flight distance but even decreasing average prices of flight distance, suggest that negative relations between distance and fare should not necessarily be considered as anomalies.

Finally, business travelers will focus more on maximizing their productivity during the trip. Therefore, they are willing to pay for a 'higher quality' service that allows last-minute bookings and changes to travel plans, and provides better check-in and on-board facilities. In addition, any ticket price increases tend to be absorbed by the firm rather than the individual traveler.

### **4.2.3 The Effect of Study Characteristics on Price Elasticities**

In addition to the economic and geographic determinants discussed in the previous section, there are a number of study characteristics that may affect the price-elasticity estimate. Also here, the availability of substitution possibilities plays a role.

The distinction between short-run and long-run elasticity estimates is important. Generally speaking, in the long-run consumers and firms are better able to adjust to price signals than in the short run, implying that the long-run demand tends to be more elastic than the short-run demand (Oum et al., 1992). In the case of transport, long run responses include geographic relocation and changes in asset holding. However, complex time-dependent behavioral patterns may induce sufficient distorting effects to prevent us from simply adopting this long-run adjustment rule as a rule of thumb. Sudden cost changes may easily lead to an exaggerated behavioral response in the short term, which may subsequently be perceived erratic and corrected for in the long run. Such behavior may in fact lead to a negative relationship between the time horizon and the price sensitivity. The actual response behavior depends on the availability of substitution possibilities. In the case of aviation demand, there are some issues worth mentioning. First, there is a lack of sufficient substitution transport modes for the air transport sector. The travel speed of the mode is still unmatched, whereas intra-modal substitution can hardly be regarded as a cost-evasive substitution due to differential fare structures and competition-related characteristics inherent in the air transport sector. Second, there are factors complicating possible long run adjustment strategies. Relocation costs, both pecuniary and non-pecuniary, as a means to evade increased transport costs tend to be relatively high due to the physical distance of the flight and the cultural distance between the departure and the arrival location.

This discussion suggests that, although short-run responses are limited to demand changes on an aggregate modal level, the use of long-run time horizons may not automatically result in higher price sensitivity estimates. The relationship between the time horizon and the

price sensitivity depends *inter alia* on the ratio of leisure travelers to business travelers, since the possibilities for demand adjustments

Most empirical studies on aviation demand have used cross-section data based on city-pairs. This offers the advantage of using larger samples than are often available in time series analysis, which is essential for studying dimensions of consumers' travel demand such as time valuation and quality of service or for studying the modal choice behavior of consumers. However, the disadvantage is that it does not always allow for accurate estimation of price and income elasticities since cross-section data generally exhibit relatively little variation in air fares per unit of distance within a given fare class (see Straszheim, 1978).

Time series data are more useful in estimating price and income elasticities, mainly because price (and income) changes have been dramatic during the last decades. However, price changes are also relatively infrequent due to government regulation.

Multicollinearity pervades both cross-section and time series estimates. In cross-section models with gravity variables, fares tend to be strongly correlated with distance variables. This problem is most severe in time series studies. Variables such as price and income tend to be tightly correlated with a time trend. Parameter estimates therefore, are generally sensitive to changes in model specification and sample coverage. Moreover, the multicollinearity, coupled with data limitations, leads to a persistent tendency to underspecify or oversimplify the model, with consequent biases in regression coefficients (see Jung and Fiji, 1976).

### 4.3 Meta-Analysis and Empirical Results

In this section, we examine the evidence offered by previous empirical case studies on the price elasticities of demand for passenger air travel by means of a meta-analysis. The objective of this section is to test whether price elasticity estimates encountered in the literature are statistically equal, and if not, to explain the variation in these elasticities in order to be able to draw some general or transferable conclusions. In Section 4.3.1, we describe the data collection process, while in Section 4.3.2 we present some descriptive statistics and variation analysis results. Subsequently, in Section 4.3.3, we discuss the results of meta-regression analysis. Finally, Section 4.3.4 ends this chapter with some concluding remarks.



### 4.3.1 Data Collection and used Variables

Since the aim of this study consists of a comparative reevaluation of previous research on price elasticities for passenger air transport we collected a number of 36 studies in which one or more price elasticity estimates were reported. This resulted in a total number of 204 observations. From each study, we collected quantitative information on the elasticity estimates and on certain geographic, economic and demographic variables and other moderator variables. An overview of the primary studies is given in Appendix 4.

The moderator variables used for explaining the variance among price elasticity estimates are transfer distance, fare class, geographic location, research method, time horizon and period of data collection. Most of the explanatory variables have been subdivided into various subcategories. Table 4.1 lists the categories for each explanatory variable.

**Table 4.1:** List of explanatory variables

Variable	Categories
Gross Domestic Product	- Continuous variable
Trend variable	- Continuous variable
Distance	- Close distance (< 500 miles) - Medium distance (500 – 1500 miles) - Long distance (>1500 miles)
Geographic scope	- North America - Europe - Australia - Intercontinental
Elasticity time horizon	- Short term elasticity - Long term elasticity
Fare class	- Economy class - Business class
Method of analysis	- Time series analysis - Cross-section analysis - Pooled analysis

GDP values are determined by taking the average GDP values for the origin and destination country of any given flight, weighted for the countries populations. These values are calculated by taking the average of the GDP values for every year of the period in which the

data, used in the original study, were collected. GDP values are corrected for price level. The trend variable is determined by calculating the average year of the period in which the data, used in the original study, were collected.

The values for the distance variable are calculated as follows. First, the average distance is calculated of all origin-destination combinations that were used to derive an elasticity estimate. Based on this average distance, the observation is categorized into one of the three distance categories as shown in Table 4.1. The geographic variable consists of the categories North America, Europe, Australia and intercontinental. The first three categories consist of intra-continental flights within the concerned continent. Flights between different continents are categorized as ‘intercontinental’.

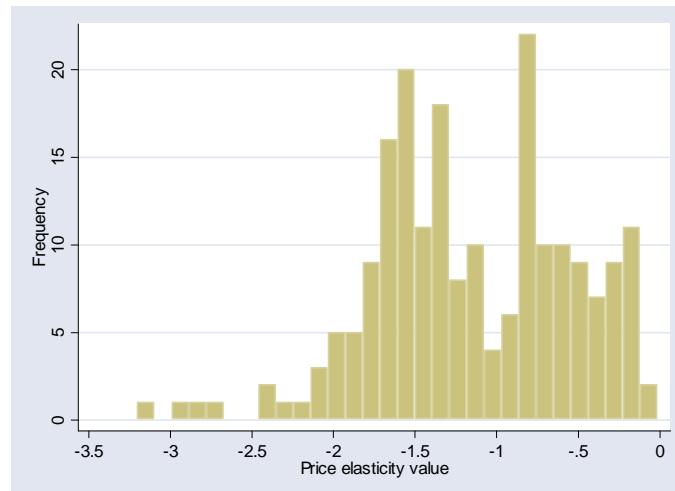
Observations are categorized as long run elasticity if the estimates are based on a dynamic specification and incorporate lagged price effects. If this is not clear from the primary study, we follow the author’s terminology. In the literature, different opinions were encountered on the definition of a long-run elasticity estimate. Despite these differences, a common underlying assumption could be discerned; a long-run estimate, as opposed to a short-run estimate, does not only signal the direct budgetary effects of a fare change on demand, but also takes into account adjustments relating to relocation and asset ownership.

### **4.3.2 Descriptive Statistics**

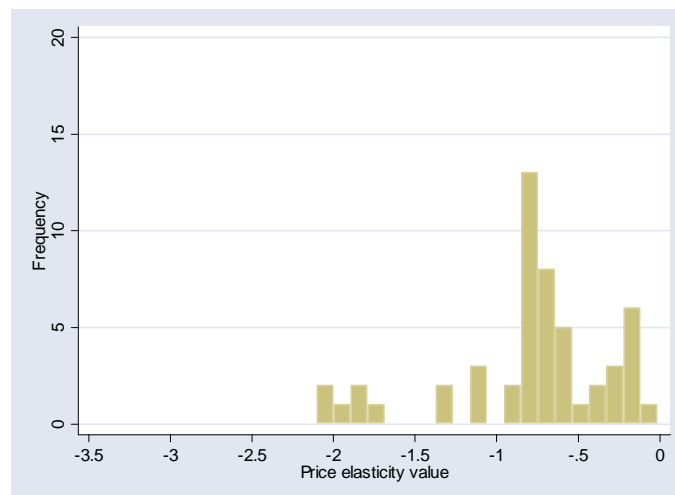
Before we proceed with meta-regression techniques to explain the variation among price elasticity estimates, we discuss some descriptive statistics of our set of estimates. Figure 4.2 shows the distribution of the total set of elasticity estimates that we collected.<sup>19</sup> The overall mean price elasticity, based on our set of 204 observations, is  $-1.146$ , which implies that a price change results in a more than proportional change in demand. The standard deviation of the elasticity distribution is  $0.619$ . The range of estimate values lies between  $-3.20$  and  $-0.01$ . The double top formation in the graph distribution is caused by the difference in price elasticity estimates between case studies that focused on business class travelers and other case studies.

---

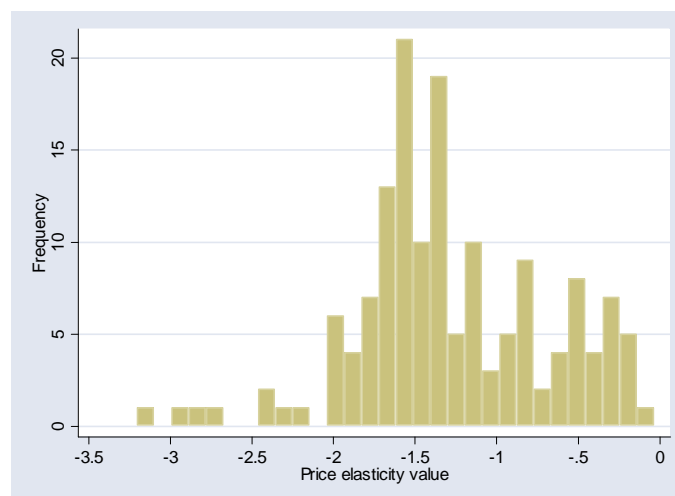
<sup>19</sup> A list of primary studies used in this analysis is given in Appendix 4.



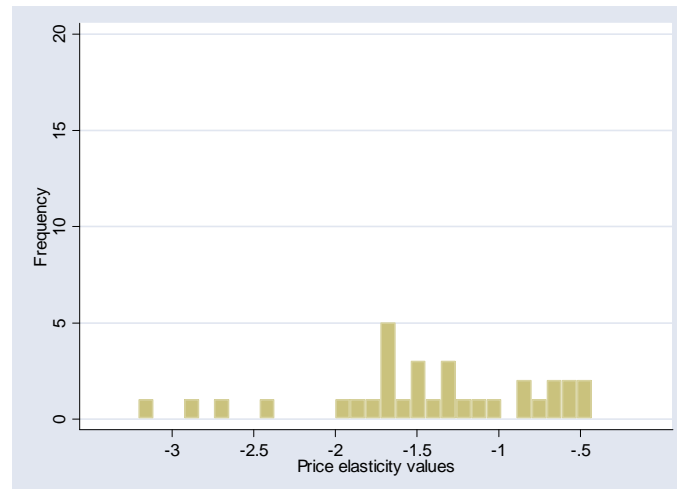
**Figure 4.2:** The distribution of price elasticity estimates of aviation demand



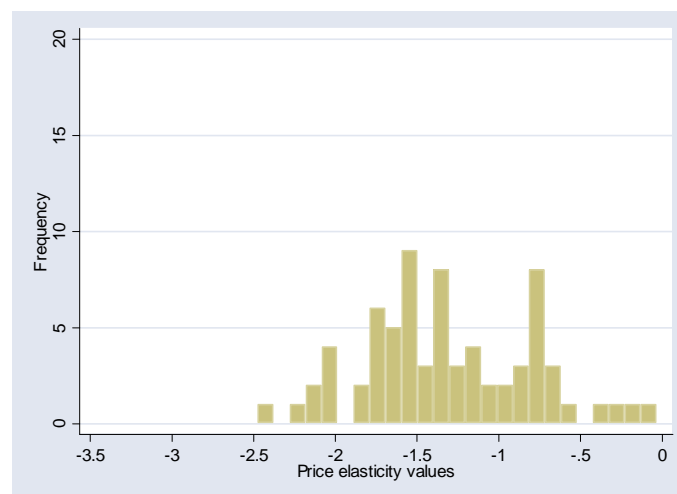
**Figure 4.3.a:** The distribution of price elasticity estimates for business class travelers



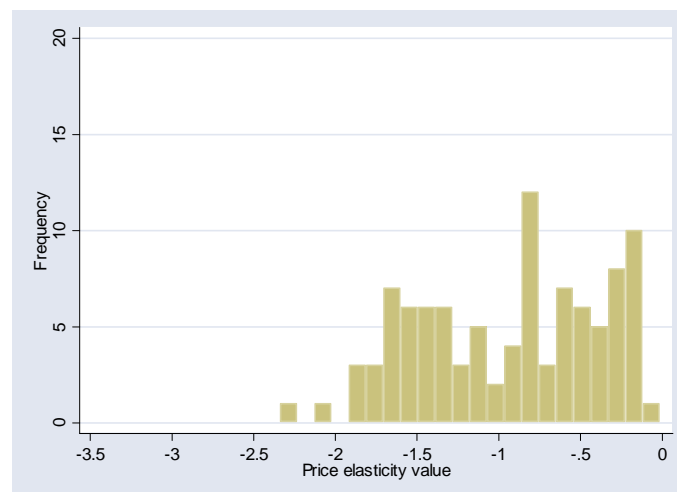
**Figure 4.3.b:** The distribution of price elasticity estimates from economy class passengers



**Figure 4.4.a:** The distribution of price elasticity estimates for short distance flights



**Figure 4.4.b:** The distribution of price elasticity estimates for medium distance flights



**Figure 4.4.c:** The distribution of price elasticity estimates for long distance flights

The distributions in Figures 4.3.a and 4.3.b show that the price sensitivity of business class passengers tends to be lower than for economy class passengers. Figure 4.4 shows the distributions of elasticity estimates for the subcategories: close distance (a), medium distance (b) and long distance (c). The relative positions of these distributions indicate that the price elasticity for long distance flights is higher than that for short distance flights. Medium distance flights tend to correspond to medium level elasticities.

**Table 4.2:** Correlation coefficients between price elasticity and various determinant factors

	Correlation coefficient	Standard error
Year of data	0.255	0.000
Distance	0.304	0.000
Intercontinental	0.466	0.000
North America	-0.350	0.000
Europe	-0.062	0.382
Australia	-0.124	0.077
Short term estimates	0.199	0.004
Long term estimates	-0.199	0.004
Economy class	-0.303	0.000
Business class	0.341	0.000
Time series studies	0.438	0.000
Pooled studies	-0.250	0.000
Cross-section studies	-0.133	0.058
GDP	-0.206	0.006

In Table 4.2, a number of correlation coefficients between the value of the price elasticity and several study characteristics are shown. The price elasticity level is positively correlated with the trend variable. Apparently, travelers have become less price sensitive over the years. Travelers are also less price sensitive, as flight distance increases. This may be due to the relative lack of substitution modes on longer distance flights. The dummy variable for intercontinental flights is positively correlated to price elasticity, which indicates the lack of substitutions for such flights. The correlation coefficient with the other geographic variables is negative. The correlation coefficient between price elasticity and short-term estimates is negative, whereas that for long-term estimates is positive. This implies that demand responses to fare price changes are higher when a longer time horizon is used. The negative correlation coefficient for the economy class dummy and the positive coefficient for the business class

dummy indicate lower price sensitivity among the latter segment. The dummy for time series studies is positively correlated with price elasticity. The dummies for cross-section studies and pooled studies show a negative correlation coefficient. Gross domestic product is negatively correlated with price elasticity, which means that travelers with a higher income tend to be more price sensitive than travelers with lower income. This may be explained from the fact that aviation transport may be regarded as a luxury good, which suggests a positive relationship between income level and the share of income share allocated to aviation transport.

### **4.3.3 Results of Meta-Regression**

In this section we perform a meta-regression analysis based on the variables shown in Table 4.1<sup>20</sup>. The dependent variable is the price elasticity of demand for passenger air transport. Since we tested positively for within-study dependence in the data set<sup>21</sup>, we use the two-level estimator given by equation (3.21). This estimator corrects for dependency between multiple observations that come from the same case study.

The positive and significant coefficient of the trend variable (see Table 4.3) indicates a negative time trend for price sensitivity. Apparently, aviation passengers have become less price sensitive over time. One could hypothesize that the process of globalization and associated growth of the aviation market have increased the dependency on air transport during the period covered by the underlying studies. This hypothesis may hold in particular for the growth of the market for long distance flights, for which there are less possibilities of substitution. The distance variable has a negative, but insignificant coefficient. Apparently, there are multiple factors exerting their influences here, thus distorting an unambiguous relationship between the distance of a flight and the price elasticity of demand for a flight. Arguably, the negative effect of distance on price sensitivity due to a relative lack of substitute modes on long distance flights may be compensated by the positive effect due to the fact that long distance flights require a larger share of the disposable income than short

---

<sup>20</sup> The variation in GDP values over time is partly captured by the time trend variable while differences in GDP over countries are partly captured by the geographic dummy variables. Since the inclusion of a GDP value, which is only available at a country level, results in multicollinearity in the dataset, we exclude it from our set of moderator variables. In addition to the variables in Table 4.1 we include a dummy variable for observations that are based on a demand model that corrects for income.

<sup>21</sup> The likelihood test, based on the difference in deviance for the least squares model and the multilevel model (see Chapter 3), is 53.76 and 51.25 for the two model specifications estimated in this section. This indicates that for both specifications significant within-study dependence is present. The within-study correlation coefficients are 0.66 and 0.63, respectively.

distance flights. Apparently, this seems to hold, as has been pointed out in Section 4.2.2, even if the (positive) correlation between fare and flight distance is lower than expected from a cost perspective. The regional dummies all enter insignificant. This is surprising, since one would expect lower price sensitivity, particularly among European passengers, due to the fact that intra-continental passenger surface transport networks offer substitutes that are not available for intercontinental flights.

In estimation 2, we exclude the dummies for North America and Australia in order to compare European price sensitivity directly with price sensitivity in North America and Australia. The quality of the European passenger surface transport network plus the fact that the European income level is low compared to North-American and Australian values would have one assume higher price sensitivity among European travelers. However, we find no significant coefficient for the Europe dummy.

**Table 4.3:** Results of meta-regression analysis

	(1)		(2)	
	Coefficient	Std. error	Coefficient	Std. error
(Constant)	-32.686 *	16.262	-27.483	17.372
Year of data	0.016 *	0.008	0.014	0.009
Distance	-0.023	0.034	-0.026	0.033
Europe	0.083	0.252	0.100	0.293
North America	0.159	0.297		
Australia	-0.342	0.261		
Long run	-0.749 **	0.093	-0.666 **	0.134
Business class	0.828 **	0.143	0.819 **	0.143
Cross-section	0.441	0.347	0.618 *	0.306
Time series	0.419	0.238	0.625 **	0.219
Dummy GDP	0.658 **	0.185	0.592 **	0.191
Between-study variance	0.43		0.39	
Within-study variance	0.22		0.23	
Within-study correlation	0.66		0.63	
N	204		204	
log-likelihood	-172.85		-172.39	

\* significant at the 5% level

\*\* significant at the 1% level

It is interesting to see that the dummies for time series and cross-section datasets are rendered significant by the exclusion of the U.S. and Australia dummies. Multicollinearity might be present, especially when for a specific region, there is a relatively large number of time series studies or cross-section studies.<sup>22</sup> The time series and cross-section dummies have the opposite sign of the Australia dummy, and the same as the (insignificant) U.S. dummy, where one would expect a relatively large number of cross-sectional studies. Note that the sign of the time series dummy is positive (and significant), which indicates that estimates from time series data yield lower price elasticities (i.e. more price sensitive passengers). This suggests that, as some authors observe, time series estimates without extensive lag structures typically yield short term elasticity estimates, although we have already included a long-term dummy<sup>23</sup>. In addition, cross-section studies yield short-term elasticities, which could explain the positive coefficient for the corresponding dummy.

The dummy coefficient for business class observations has a positive sign and is significant. Business class travelers' price elasticities are lower on average than economy class travelers' price elasticities. The lower price sensitivity in the business class segment can be explained by the higher valuation of time among these travelers. Other explanations are that (i) for business class flights the number of substitutes is lower than for leisure flights, (ii) business class travelers are more willing to pay for 'higher quality' services (such as allowance for last-minute bookings and changes in travel plans, better check-in and on-board facilities) and (iii) tickets are usually paid by the firm, rather than the individual traveler. While interpreting these results, it is important to note that travelers with a business motive do not necessarily travel by business class. To a lesser degree, this also holds for leisure travelers.

The coefficient of the dummy for long-run elasticities is negative, which implies that a longer response time enables consumers to adjust better to changes in fare prices. Even if such long-run demand adjustment effects are expected to be less strong in aviation transport than in other modes, the results show that a long-run adjustment time decreases the price elasticity. Finally, the coefficient for the dummy that corrects for the inclusion of an income related variable in the demand equation is positive and significant. Apparently, if the results

---

<sup>22</sup> Correlation results indeed indicate a strong correlation between the time series dummy and intercontinental flights and, to a lesser extent, between the time series dummy and the North America dummy.

<sup>23</sup> In a plot of a sufficiently long time series, convergence to equilibrium might be visible. However, a simple regression (of traffic on fare) would weight the data point representing the initial impact of the fare change the same as the data point representing the long run equilibrium. The estimated fare coefficient is therefore biased towards the initial impact. Because the cross-sectional variance is larger, the estimation results from a cross-sectional study are therefore more likely to be indicative of the long-run effect (see Abrahams, 1983).



of the original study are corrected for income level, this leads to higher price elasticities. In theory, omission of the income variable in the demand equation causes the price parameter to be biased. When price and income are the only arguments in the demand equation and if the income is falsely omitted, the bias is given by the income parameter multiplied by  $\text{Cov}(\text{price}, \text{income}) / \text{Var}(\text{price})$  (see for example, Greene, 1992). When this covariance is negative (for example, because in the course of time income increases in real terms and fares decrease in real terms as a result of deregulation), the bias has a negative sign and the estimated price sensitivity is too high. Thus, inclusion of the income would yield lower price sensitivity.

## 4.4 Conclusion

When formulating an environmental policy concerning aviation, the authorities (also) need information on the price sensitivity of passengers to predict the effectiveness of the policy. For example, if the airlines or airports can charge all extra costs to the passengers without decreasing demand, the policy has no other effect than increasing the authorities' revenues. In this study, price elasticities of demand for passenger air transport are analyzed. This research synthesis provides valuable information on the "expected" degree of price sensitivity in a specific setting. It can also provide information on the design of new research or questionnaires.

The results of the meta-analysis show that, with a mean price elasticity of -1.15, the demand for aviation transport is elastic, which indicates that an increase in price will lead to a more than proportionate decrease in demand on average. This suggests that fuel charging is an effective instrument to decrease demand and to direct airlines towards more environmentally friendly policies. From the meta-regression analysis, the following results are obtained. Long-run price elasticities are higher in absolute value, as expected from a theoretical point of view. Hence, basing long run policy instruments on short run elasticities leads to distortions. Passengers become more price sensitive over time; this also needs to be acknowledged in the design of long run policy instruments. Surprisingly, we find that European passengers are not more price sensitive than U.S. passengers and Australian passengers, while one would expect that the availability of more substitutes would result in higher price sensitivity within Europe. Omission of the income variable in the underlying study decreases the price elasticity. When the results of the study are to be used, it is important that one is aware of this bias. This finding also suggests that in setting up a new study, income should not be left out. Finally,

business passengers are less sensitive to price; this is a common finding in the literature. The difference is about 0.6, *ceteris paribus*. This fact gives the airlines the opportunity to charge part of the extra costs (resulting from a price-based policy instrument) to the passengers, where the business passengers can be charged more than proportional without decreasing demand. If this is not acknowledged by the authority, the (environmental) policy may have little success, although the authority gets the revenues.

In general, the results suggest that price charging is an effective instrument to decrease demand and to promote environment-friendly policies; additionally, the revenues from the charges can be used to finance a charging system that combines price charges and subsidization of environment-friendly aviation technologies. The development and application of such charging systems is supported by current EU policy on domestic flights charges and EU proposals on aviation fuel tax exemptions and promotion of environment-friendly technologies and operations

The research agenda that follows from this chapter is as follows. First, attention should be paid to the question whether research synthesis allows for the determination of confidence intervals for price elasticities in specific cases. In order to do so, one needs accurate data on a wide range of (study) characteristics. For example, the finding on price sensitivity of European passengers necessitates additional analysis, using more and more specific characteristics. Second, the application of research synthesis, although a very useful technique, does not remove the need for specific case studies. Research synthesis may be necessary because performing new case studies may be too expensive or time-consuming, but it can only be carried out based on a sufficient number of primary studies.

## **Appendix 4: List of primary studies on the price elasticity of aviation demand**

- |                                    |                                  |
|------------------------------------|----------------------------------|
| Abrahams (1983)                    | Marin (1995)                     |
| Agarwal and Talley (1984)          | Melville (1997)                  |
| Alperovich and Machnes (1994)      | Milloy et al. (1985)             |
| Andrikopoulos and Terovitis (1982) | Morrison and Thompson (1986)     |
| Brown and Watkins (1974)           | Morrison and Thompson (1989)     |
| BTCE (1983)                        | Mutti and Murai (1977)           |
| BTCE (1986)                        | Okeahialam (1990)                |
| BTCE (1988)                        | Oum and Gillen (1982)            |
| BTCE (1995)                        | Oum et al. (1986)                |
| De Vaney (1974)                    | Oum et al. (1993)                |
| Fridstrom and Thune-Larsen (1989)  | Smith and Toms (1978)            |
| Haitovsky et al. (1987)            | Straszheim (1978)                |
| Hartley and Trengrove (1991)       | Talley and Schwarz-Miller (1987) |
| Hensher and Louviere (1983)        | Taplin (1980)                    |
| Ippolito (1981)                    | Taplin (1997)                    |
| Jud and Hyman (1974)               | Thompson and Caves (1992)        |
| Jung and Fuji (1976)               | Wheatcroft (1956)                |
| Liew and Liew (1979)               |                                  |

## **Chapter 5**

# **A Meta-Analysis of Price Elasticities of Gasoline Demand. A System of Equations Approach**

### **5.1 Introduction**

In a world where energy-conservation is a key issue for policymakers, the demand for gasoline is a highly interesting topic, both from a political, environmental and economic viewpoint. Hence, it also receives much attention as a research topic. Understanding the determinants of gasoline demand has been of interest to economists for almost three decades. Especially since the 1973 energy crisis, there have been a growing number of study efforts to model the demand for gasoline. Initially, studies mainly addressed concerns about the availability of depletable resources and national security concerns raised by the oil supply shocks of the 1970s. Lately, studies increasingly addressed the various environmental consequences of gasoline consumption, particularly with respect to the emission of greenhouse gases (Kayser, 2000). In these studies on gasoline demand, particular attention has always been paid to the impact of the price level of gasoline. For environmental and political reasons, policy makers are highly interested in the impact of gasoline taxes and autonomous price changes on the demand for gasoline; in order to assess the effectiveness of gasoline pricing policies, information about the relationship between gasoline prices and demand is required. In this context, it is important to note that the total demand for gasoline is a composite function of the variables fuel efficiency, car use (the number of kilometers per car) and car ownership. A change in gasoline price affects each of these variables. The size of these partial effects determines the total impact of a price change on the demand for gasoline.

Empirical estimates of the price elasticity of gasoline vary a lot between studies. This is hardly surprising, given the large amount of variation in estimation techniques, model specifications and other study characteristics. Hence, the use of meta-analysis appears to be a very useful approach to gain more insights into the price elasticity of gasoline demand. In this chapter, we aim to (i) estimate the mean value for the price elasticity of gasoline demand and

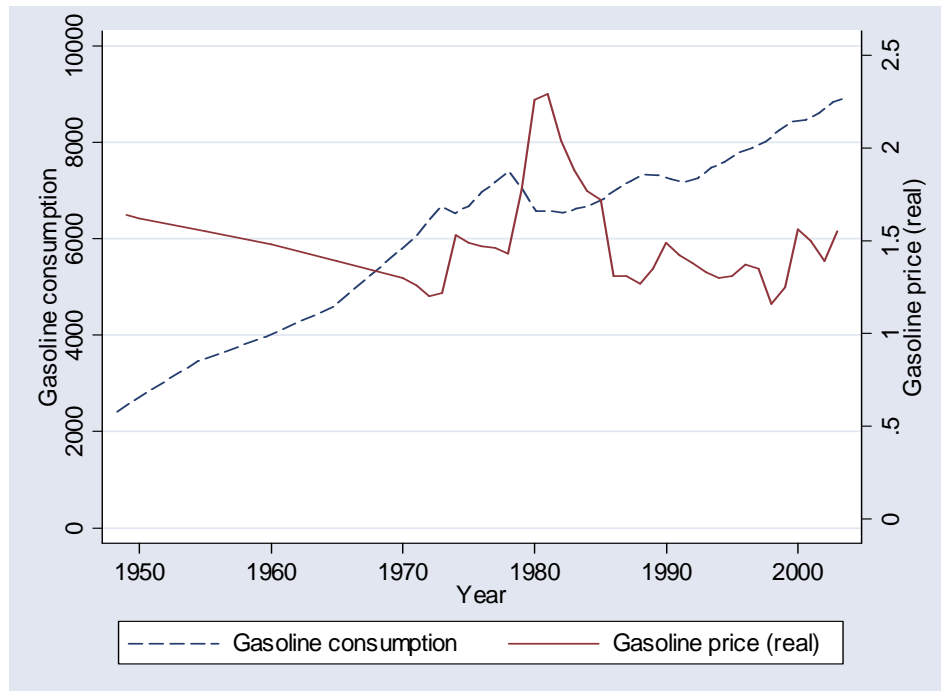
to decompose this into mean values for the elasticities of fuel efficiency, car use and car ownership with respect to gasoline price and (ii) determine the impact of study characteristics on the estimated elasticity values.

In order to do so, we collect observations on price elasticities of gasoline demand as well as observations on price elasticities of fuel efficiency, car use, car ownership, traffic volume and gasoline consumption per car. We use a system of equations approach, in which we utilize the functional relationship between the elasticities in order to extract as much information as possible from the literature. In contrast with conventional techniques used in earlier meta-analytical studies on the price elasticity of gasoline demand (Espey, 1998; and Graham and Glaister, 2002; Hanly, 2002), this approach enables us to combine information of different types of elasticities in order to obtain more accurate estimates. Furthermore, it enables us to interpret and explain the results in a more precise and detailed fashion.

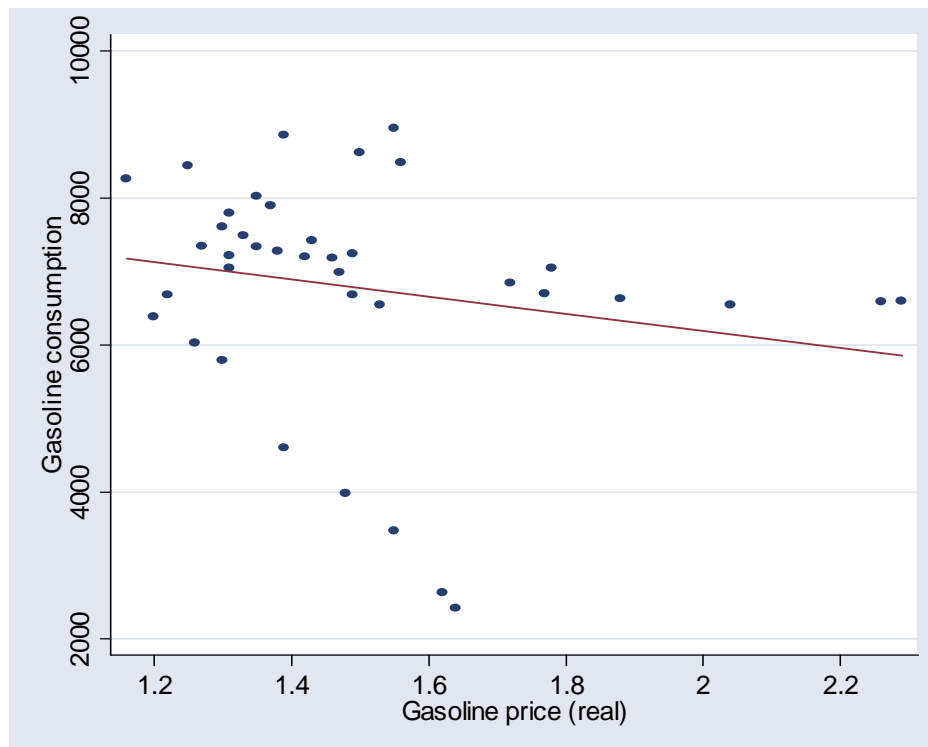
This chapter is structured as follows. In Section 5.2 we discuss how the price elasticity of demand for gasoline can be expressed as a linear function of the price elasticities of fuel efficiency, car use and car ownership. Section 5.3 introduces a system of equations method that, within the framework of a meta-analytical study, integrates and combines heterogeneous but conceptually related effect sizes by utilizing the linear relationship between them. In Section 5.4, we present the combined mean elasticity values that we estimated. Section 5.5 shows the results of a conventional meta-regression analysis and those of the system of equations approach that we discuss in Section 5.3. Section 5.6 concludes.

## **5.2 Price Sensitivity in the Context of Gasoline Demand**

Figure 5.1 shows the development of the real price of gasoline and the consumption of gasoline in the US - the largest consumer of gasoline - between 1949 and 2003. During this period gasoline consumption has grown from 2,410 to 8,937 thousand barrels per day. Growth was particularly rapid in the period between 1949 and 1973. Between 1973 and 1992, the growth rate slowed down and even reached negative values in certain years. After 1992, the growth rate has been strictly positive again, although in a less dramatic fashion than in the period before 1973. With respect to gasoline price there has been considerable variation between 1949 and 2003, with a minimal value of \$1.16 per liter in 1998 and a maximum value of \$2.29 per liter in 1981. First, during the period between 1949 and 1973 the gasoline price gradually decreased.



**Figure 5.1:** Real gasoline price and annual gasoline consumption between 1949 and 2003. Gasoline consumption is in thousands barrels per day. Gasoline price is in dollars per liter. Source: The Annual Energy Review (2003).



**Figure 5.2:** The relation between real gasoline price and annual gasoline demand in the period between 1949 and 2003. Source: based on historical data from the Annual Energy Review (2003).

As a consequence of three major political events, i.e. the energy crisis caused by the OPEC oil embargo (1973), the Iranian revolution (1979) and the Iraq-Iran war (1980-1988), the gasoline price was very high during the period between 1973 and 1985. Since 1985, the gasoline price has been relatively stable again. The scatter plot in Figure 5.2 shows that the correlation between gasoline consumption and gasoline prices has been negative, with a correlation coefficient value of -0.50.

Total gasoline demand can be decomposed in a number of components as follows:

$$\text{Gasoline demand} = \frac{\text{Gasoline demand}}{\text{Total car kms}} \times \frac{\text{Total car kms}}{\text{Number of cars}} \times \text{Number of cars} \quad (5.1)$$

We can rewrite this as:

$$\text{Gasoline demand} = \text{Efficiency}^{-1} \times \text{Car use} \times \text{Car ownership} \quad (5.2)$$

Changes in gasoline prices affect each of these components to a certain degree. The relative size of those partial effects depends on the behavior response of the consumer when faced with a gasoline price change<sup>24</sup>. For example, in the case of a price increase, a consumer may decide to travel less by car, either by making more use of other transport modes or by traveling less in general. Furthermore, a car owner may decide to sell her car or switch to a more fuel-efficient type. All of these responses ultimately affect the consumption of gasoline. Thus, a change in gasoline price affects the total demand for gasoline via (i) fuel efficiency, (ii) car use and (iii) car ownership. The response of consumers to a change in price may be different in the short run and in the long run. In the short run, people might respond to a change in price primarily by changing their car usage. Switches to more fuel-efficient cars and changes in the number of cars are probably of more importance in the long run.

The sensitivity of total gasoline demand to changes in the gasoline price is measured by calculating or estimating the price elasticity of demand,  $\varepsilon_G$ . This indicator measures the responsiveness of quantity demanded to a change in price, with all other factors held constant. It is defined as the magnitude of a proportionate change in quantity demanded divided by a

---

<sup>24</sup> While the impact of gasoline price on car use and car ownership depends largely on consumer behavior, the impact of gasoline price on *fuel efficiency* is, to a certain extent, determined by fuel efficiency target policy such as the 1975 Energy Policy and Conservation act (see for example Greene 1990).

proportionate change in price. If the changes are taken infinitely small, the term *point elasticity* is used. The (point-) price elasticity of demand is expressed as<sup>25</sup>

$$\varepsilon_G = \frac{\partial \text{Demand}}{\partial \text{Price}} \frac{\text{Price}}{\text{Demand}} = \frac{\partial \ln \text{Demand}}{\partial \ln \text{Price}} \quad (5.3)$$

By substituting (5.2) in (5.3),  $\varepsilon_G$  can be decomposed as follows:

$$\varepsilon_G = -\varepsilon_{FE} + \varepsilon_{KM/C} + \varepsilon_C \quad (5.4)$$

Where  $\varepsilon_G$ ,  $\varepsilon_{FE}$ ,  $\varepsilon_{KM/C}$  and  $\varepsilon_C$  represent the elasticities of gasoline demand, fuel efficiency, car use and car ownership with respect to gasoline price, respectively. These elasticities indicate the response in gasoline demand, fuel efficiency, car use and car ownership to a change in the price of gasoline. Note that there is a linear relationship between the elasticities. By subtracting  $\varepsilon_C$  on both sides of (5.4) establish a linear relationship between the price elasticity of gasoline consumption per car and the price elasticities of fuel efficiency and car use:

$$\varepsilon_{G/C} = -\varepsilon_{FE} + \varepsilon_{KM/C} \quad (5.5)$$

The price sensitivity of fuel demand per car is thus decomposed into sensitivity measures for fuel efficiency and car use. Finally, by adding  $\varepsilon_{FE}$  to (5.4) we establish a linear relationship between the price elasticity of traffic volume (total number of car kilometers) and the price elasticities of car use and car ownership:

$$\varepsilon_{KM} = \varepsilon_{KM/C} + \varepsilon_C \quad (5.6)$$

The price elasticity of traffic volume is decomposed into price sensitivity measures for car use and car ownership. In the next section, we develop a meta-analysis estimation model that is based on a system of meta-regression equations. These meta-regression equations are based on the relationships that we established in (5.4), (5.5) and (5.6).

---

<sup>25</sup> Price elasticities of demand are typically estimated as a double log equation:  $\ln \text{Demand} = \alpha \ln \text{Price} + \mathbf{X}\beta + \mu$ . The coefficient  $\alpha$  is used as an estimate for  $\varepsilon_G$ .  $\mathbf{X}$  and  $\beta$  represent the data matrix and coefficients with respect to other explanatory variables.



### 5.3 A Model based on a System of Meta-Regression Equations: the SMR Model

In the introduction to this chapter, we mentioned that due to the large variation between empirical estimates of the price elasticity of gasoline, the use of meta-analysis seems to be a very useful approach to gain insight into the mean value of the price elasticity and into the effect of conditioning factors on the elasticity value. Previous studies such as Espey (1998), Hanly et al. (2002) and Graham and Glaister (2002) have used such an approach to investigate the price elasticity of gasoline demand. These studies investigate the variation in the elasticity estimates and use a linear regression approach based on the following general model:

$$\varepsilon_G = \mathbf{X}\beta + \mu \quad (5.7)$$

where  $X$  denotes the set of moderator variables and  $\mu$  denote the disturbance term associated with the sampling error of the price elasticities. While some of the mentioned studies discuss the relationship between the price elasticities of gasoline demand, fuel efficiency, car ownership and car use, their meta-analyses focus exclusively on the price elasticities of gasoline demand.

In the literature, estimates are found for each of the six elasticities discussed in Section 5.2, together with information about the studies they come from. From a meta-analytical perspective, this raises the interesting question if the linear relationship between the different elasticities can be exploited in order to combine all the estimates in a meta-analysis. In the remainder of this section, we investigate these questions in more detail. Based on the various types of elasticity estimates that are found in the literature and the information we have with respect to the studies they come from we can formulate the following set of meta-regression models:

$$\varepsilon_G = \mathbf{X}_G\beta_G + \mu_G \quad (5.8)$$

$$\varepsilon_{FE} = \mathbf{X}_{FE}\beta_{FE} + \mu_{FE} \quad (5.9)$$

$$\varepsilon_{KM/C} = \mathbf{X}_{KM/C}\beta_{KM/C} + \mu_{KM/C} \quad (5.10)$$

$$\varepsilon_C = \mathbf{X}_C\beta_C + \mu_C \quad (5.11)$$

$$\epsilon_{G/C} = \mathbf{X}_{G/C} \beta_{G/C} + \mu_{G/C} \quad (5.12)$$

$$\epsilon_{KM} = \mathbf{X}_{KM} \beta_{KM} + \mu_{KM} \quad (5.13)$$

By substituting (5.9), (5.10) and (5.11) into (5.4), (5.5) and (5.6) we obtain the following equations:

$$\epsilon_G = -\mathbf{X}_{FE} \beta_{FE} + \mathbf{X}_{KM/C} \beta_{KM/C} + \mathbf{X}_C \beta_C + \mu_G \quad (5.14)$$

$$\epsilon_{G/C} = -\mathbf{X}_{FE} \beta_{FE} + \mathbf{X}_{KM/C} \beta_{KM/C} + \mu_{G/C} \quad (5.15)$$

$$\epsilon_{KM} = \mathbf{X}_{KM/C} \beta_{KM/C} + \mathbf{X}_C \beta_C + \mu_{KM} \quad (5.16)$$

Where  $\mu_G = -\mu_{FE} + \mu_{KM/C} + \mu_C$ ,  $\mu_{G/C} = -\mu_{FE} + \mu_{KM/C}$  and  $\mu_{KM} = \mu_{KM/C} + \mu_C$ . For estimates of  $\epsilon_G$ ,  $\epsilon_{G/C}$  and  $\epsilon_{KM}$ , collected from the literature we have by construction that  $\mathbf{X}_G = \mathbf{X}_{FE} = \mathbf{X}_{KM/C} = \mathbf{X}_C$ <sup>26</sup> Therefore, we can rewrite (5.14), (5.15) and (5.16) as

$$\epsilon_G = \mathbf{X}_G (-\beta_{FE} + \beta_{KM/C} + \beta_C) + \mu_G \quad (5.17)$$

$$\epsilon_{G/C} = \mathbf{X}_{G/C} (-\beta_{FE} + \beta_{KM/C}) + \mu_{G/C} \quad (5.18)$$

$$\epsilon_{KM} = \mathbf{X}_{KM} (\beta_{KM/C} + \beta_C) + \mu_{KM} \quad (5.19)$$

Note that by comparing (5.17), (5.18) and (5.19) with (5.8), (5.12) and (5.13) we see that we have the following equalities with respect to the estimation coefficients:  $\beta_G = -\beta_{FE} + \beta_{KM/C} + \beta_C$ ,  $\beta_{G/C} = -\beta_{FE} + \beta_{KM/C}$  and  $\beta_{KM} = \beta_{KM/C} + \beta_C$ . The models in (5.17), (5.18) and (5.19) are estimated simultaneously with (5.9), (5.10) and (5.11) by solving for the vector coefficients  $\beta_{FE}$ ,  $\beta_{KM/C}$  and  $\beta_C$  in order to minimize the total sum of squared residuals of all equations.

Thus, by utilizing the linear relationship between  $\epsilon_G$ ,  $\epsilon_{G/C}$ ,  $\epsilon_{KM}$ ,  $\epsilon_{FE}$ ,  $\epsilon_{KM/C}$  and  $\epsilon_C$  we have constructed a model based on a system of meta-regression equations, which allows us to combine estimates of different types of elasticities so that we can use more information than in a conventional meta-analysis. This enables us to (i) estimate the overall price elasticity of gasoline demand and decompose this into price elasticities of fuel efficiency, car use and car

<sup>26</sup> For this to hold, it is required that the sets of moderator variables associated with estimates of  $\epsilon_G$ ,  $\epsilon_{KM}$ ,  $\epsilon_{G/C}$ ,  $\epsilon_{FE}$ ,  $\epsilon_{KM}$  and  $\epsilon_C$  are chosen to be identical. Note that the choice of the moderator set is under the control of the analyst.

ownership and (ii) estimate meta-regression results with respect to the price elasticity of gasoline demand and decompose them into results with respect to the other price elasticities.

## 5.4 Constrained Estimation of Mean Effect Sizes

Our meta-analysis is based on a set of 312 observations of elasticity estimates, coming from 43 primary studies<sup>27</sup>. In this set, we have 158 price elasticities of total gasoline demand ( $\epsilon_G$ ), 111 price elasticities of gasoline demand per car ( $\epsilon_{D/C}$ ), 10 price elasticities of traffic volume ( $\epsilon_{KM}$ ), 15 price elasticities of fuel efficiency ( $\epsilon_{FE}$ ), 3 price elasticities of car use ( $\epsilon_{KM/C}$ ) and 15 price elasticities of car ownership ( $\epsilon_C$ ). Figure 5.3 shows the distribution of the observations of price elasticities of total gasoline demand from our dataset. While the total range of elasticity values lies between -2.04 to 0.28, the vast majority of the elasticities lie between -1.0 and zero. Furthermore, we see that the distribution is skewed in the sense that the value zero appears to be a cut-off point. Based on the results of a  $Q$ -test<sup>28</sup> the hypothesis of homogeneity should be rejected. Hence, in Section 5.5 we investigate the variation in the effect sizes in a multivariate way.

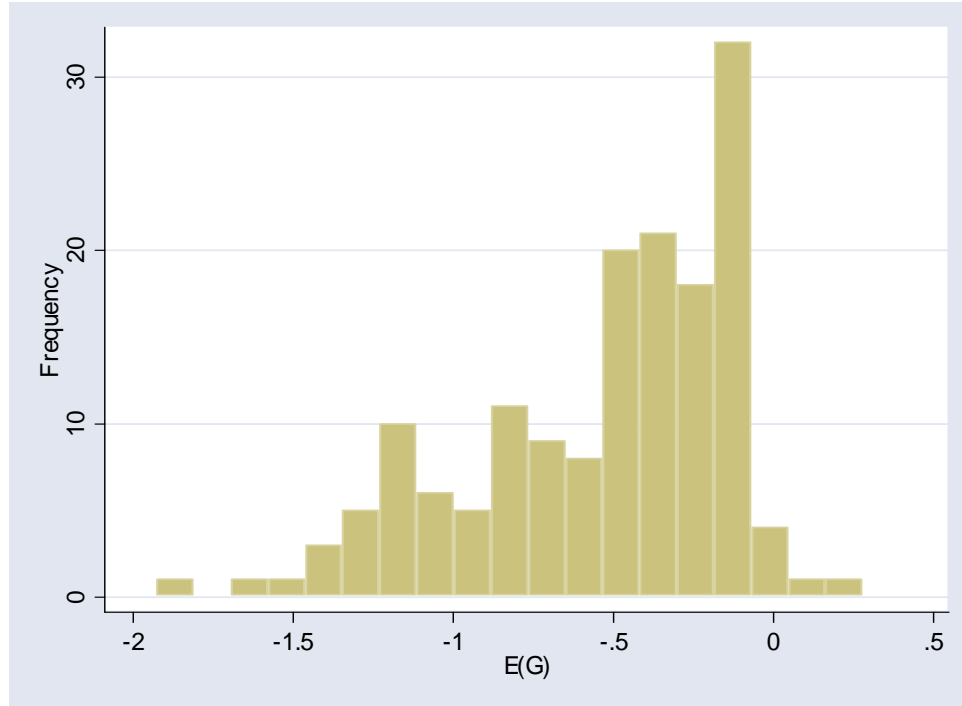
Table 5.1 shows the simple means, constrained means and some summary statistics of all sets of elasticity estimates. The simple means are calculated with the standard errors of the estimates used as weights. Some of the calculated means have unexpected values. For example,  $\epsilon_{KM/C}$ , the mean price elasticity of car use, which is one of the components of total gasoline demand is larger than  $\epsilon_G$ , the price elasticity of total gasoline demand. Apparently, the number of observations for some of the elasticities is too low to calculate reliable means.

This problem can be resolved by using constrained means. These are calculated by simultaneously regressing each elasticity type on a vector of ones, with (5.4), (5.5) and (5.6) used as constraints, and the standard errors of the estimates as weights. This approach utilizes the relationship between the elasticities to combine all information there is available and leads to more realistic mean elasticity values. The results show that the mean price elasticity of gasoline demand is -0.53. This value lies between the values found in Espey (1998), Hanly et al. (2002) and Graham and Glaister (2002), which are -0.44, -0.69 and -0.45, respectively (see Table 5.2). Apparently, automobilists are not very sensitive to gasoline price changes. The

<sup>27</sup> A list of primary studies used in this analysis is given in Appendix 5.

<sup>28</sup> The value of the  $Q$ -statistic is 1717.927 on 157 degrees of freedom. Based on this result, the hypothesis of homogeneity should be rejected.

price elasticity of gasoline demand (-0.53) can be decomposed into the price elasticities of fuel efficiency (0.22), car use (-0.10) and car ownership (-0.22). These results indicate that the response in demand resulting from a change in gasoline price is mostly driven by responses in fuel efficiency and car ownership and to a lesser degree by a response in car use.



**Figure 5.3:** The distribution of price elasticity estimates of gasoline demand

**Table 5.1:** Descriptive statistics of price elasticity estimates of fuel efficiency, car use, car ownership, gasoline consumption per car and traffic volume

Stats	$\epsilon_G$	$\epsilon_{FE}$	$\epsilon_{KM/C}$	$\epsilon_C$	$\epsilon_{G/C}$	$\epsilon_{KM}$
N	158	15	3	14	111	10
Simple means	-0.54	0.12	-0.55	-0.32	-0.32	-0.04
Constrained mean	-0.53 **	0.22 **	-0.10	-0.22 **	-0.32 **	-0.32 **
Std. error	0.04	0.02	0.03	0.15	0.03	0.13
Min	-2.04	0.00	-0.61	-0.70	-1.21	-0.33
Max	0.28	0.29	-0.50	0.16	0.37	1.14

\*\* significant at the 1% level

In Hanly et al. (2002) higher absolute values are found for the price elasticities of car use and car ownership (see Table 5.2). However, these elasticities are estimated separately without using the linear relationship between them. As a result, they are based on a relatively small number of estimates. Furthermore, the table shows that the absolute value of the mean price elasticity of fuel efficiency found in Graham and Glaister (2002) is higher than the value we find. Graham and Glaister make use of the relationship between the elasticities but take a

different approach. They estimate the price elasticities of gasoline demand and traffic volume separately and use these estimates to calculate the elasticity of fuel efficiency.

**Table 5.2:** Descriptive statistics of price elasticity estimates of fuel efficiency, car use and car ownership<sup>29</sup>

Stats	$\epsilon_G$	$\epsilon_{FE}$	$\epsilon_{KM/C}$	$\epsilon_C$	$\epsilon_{G/C}$	$\epsilon_{KM}$
Graham and Glaister (2002)	-0.688 **	0.373 **	-	-	-	-0.312 **
Hanly et al. (2002)	-0.450 *	-	-0.303 *	-0.148	-0.324	-0.257 *
Espey (1998)	-0.442 **	-	-	-	-	-

\*significant at the 5% level

\*\*significant at the 1% level

In general, the inelastic results that we find suggest that automobilists are not very sensitive to changes in fuel prices. Hence, pricing policy may not be a very effective instrument to decrease the demand for gasoline.

## 5.5 Results of Meta-Regression Analysis

In the previous section, we tested for heterogeneity in the effect sizes. The test result indicated that such heterogeneity is indeed present. In this section, we investigate the variation in elasticity estimates in a multivariate way. First, we carry out a conventional meta-regression analysis based on the model in (5.8), in order to investigate the impact of study characteristics on the estimated value of the price elasticity of gasoline demand. Next, we estimate the system of meta-regression equations model (SME) discussed in Section 5.4 in order to obtain results that are enhanced with the information we have about estimates of the other types of elasticities and the studies they come from. Furthermore, estimating the SME model will enable us to investigate the impact of study characteristics on the estimated value of the other types of elasticities. In order to account for observation (in)accuracy we use a weighted estimation procedure, with the weights based on the inverse of the standard errors<sup>30</sup>.

<sup>29</sup> The reported mean values are calculated by taking the average of the mean short- and long-run estimates reported in these studies, while weighting for the number of observations. This also holds for the values of the partial elasticities.

<sup>30</sup> The likelihood test based on the difference in deviance for the least squares model and the multilevel model (see Section 4.6) is 5.07 for the conventional meta-regression and 3.48 for the integrated model. This indicates the presence of within-study dependence in the conventional model (the within-study correlation coefficient is 0.38) and the absence of within-study dependence in the integrated model. As we focus primarily on the integrated meta-regression model and we use the conventional model mainly for comparison we use weighted least squares estimation for both models.

Table 5.3 shows the list of study characteristics that we include as moderator variables in order to investigate the variation in the effect size. Most of these are categorical variables. To identify and correct for regional differences in price elasticity estimates we use a dummy variable to distinguish estimates based on US, Canada or Australia (UCA) data from estimates that are based on data from other countries. We include two temporal variables. First, we use a dummy variable for studies based on data from the period between 1974 and 1981, to correct for a change in impact during the period following the oil crisis. Furthermore, we include the average year of data sources as a continuous variable. This allows us to investigate if there is a time trend in the observed estimates.

**Table 5.3:** List of explanatory variables

Name of variable	Categories
Geographical region	(UCA, Other regions)
Trend variable	Continuous variable
Data period	(1974 - 1981, other period)
Database type	(Cross-section, Time series, Pooled CS-TS)
Time horizon	(Short run, Long run)
Dynamics	(Dynamic model, Static model)
Functional form	(Loglinear, Non-linear)
Number of explanatory variables	Continuous variable
Included variables	(Lagged price, No lagged price)

Next, we include dummy variables for the type of database that is used in the primary studies; we use dummies to investigate differences in effect size between time series data, cross-country data, pooled time series and cross-section data.

With respect to the functional form of the demand specification, we distinguish between dynamic and static models. We categorize an observation as dynamic if the model includes a lagged dependent variable or a (series of) lagged gasoline price(s). In order to assess the impact of response time on price sensitivity, we include a dummy variable to distinguish between short run and long run elasticities. Observations are categorized as long run elasticity if the estimates are based on a dynamic specification and incorporate lagged price effects. Furthermore, we use a dummy variable to distinguish between loglinear and nonlinear models. If an explanatory variable that is correlated with gasoline price is excluded from the gasoline demand equation then the effect of this variable on demand is (partly) picked up by the coefficient of the price variable. The price elasticity estimate is then biased.

Important explanatory variables that are typically included are variables related to income, car ownership and prices of other commodities. Exclusion of any of these variables could affect price elasticity estimates. We include a moderator variable for the number of explanatory variables included in the demand equation in the primary study.

**Table 5.4:** Estimation results from a standard meta-regression analysis

Variable	Coefficient	Std. error
(Constant)	9.879	9.975
UCA	0.130 *	0.064
Data Year	-0.005	0.005
Oil_crisis	0.002	0.113
Cross-section	-0.439 **	0.099
Pooled	-0.031	0.067
Long-run	-0.458 **	0.069
Dynamic	0.160 **	0.070
Non-linear	-0.229 *	0.120
Included variables	0.022 *	0.010
N	158	
R <sup>2</sup> -adjusted	0.387	

\* significant at the 5% level

\*\* significant at the 1% level

Table 5.4 shows the result of the conventional meta-regression analysis. The dependent variable is the price elasticity of total gasoline demand. The significantly negative coefficient of the regional dummy indicates that price sensitivity is lower in the US, Canada and Australia. Espey (1998) and Hanly et al (2002) found similar results. Two explanations could be offered. First, the combination of high income and low gasoline prices makes consumers less price sensitive. Second, car dependence is higher in these countries, due to the combination of sparse population and relatively underdeveloped public transport infrastructure.

The coefficient of the trend variable shows that there is no significant time trend in the price elasticity of gasoline demand. Furthermore, the oil crisis coefficient indicates that during the years after the oil embargo there is no significant change in price sensitivity. The negative coefficient of the dummy for cross-section studies may be related to the fact that, as some authors observe, time series estimates without extensive lag structures typically yield

short term elasticity estimates while the estimation results from a cross-sectional study are more likely to be indicative of the long-run effect (see Abrahams, 1983).

Long-run price sensitivity appears to be significantly stronger than short-run price sensitivity. This is a common finding in the literature, including the studies by Espey (1998), Graham and Glaister (2002) and Hanly et al. (2002). The reason for this is that the longer time period gives consumers more possibilities to respond to the price change. If we look at the functional form of the demand equation, we see that a dynamic model significantly decreases the price sensitivity. Similar results were found in Espey (1998) and Graham and Glaister (2002). Because of the negative correlation between price and lagged demand and the positive correlation between demand and lagged demand, this result is in agreement with the omitted variable formula (see Greene, 2003, p.148). Use of a non-linear demand model does not have a significant impact on the estimated price elasticities. This seems to indicate that price sensitivity behavior can be adequately modeled with a (log-)linear equation. Finally, the number of included variables decreases the price sensitivity. This indicates that the use of a small set of explanatory variables in the demand specification increases the probability of omitted variable bias.

Table 5.5 shows the estimation results of the SMR model. Column 1 shows the impact of study characteristics on the price elasticity of total gasoline demand. Column 2, Column 3 and Column 4 show the impact of study characteristics on the price elasticities of fuel efficiency, car use and car ownership, respectively. Comparing the results in Column 1 with the results of the conventional meta-regression analysis in Table 5.5 shows that the changes in coefficient values are small, despite the fact that 153 additional observations are included. This indicates robustness of the results. The sign pattern also seems robust, with only one (insignificant) coefficient changing signs. There are some changes in the significance. For the systems of equations model, the standard errors of all coefficients are smaller. This is probably caused by an increase in the sample size as a result of the inclusion of extra observations. The coefficients of the dummies for UCA studies, cross-section studies, long run studies and dynamic specification, which are significant when using the unconstrained meta-regression estimation, remain significant when using the SMR model. In addition to these, the time trend variable enters significantly under the SMR model. The coefficient of the number of included variables becomes insignificant. The negative coefficient of the trend variable indicates that consumers have become more price sensitive over the years. This could be due to the increase in gasoline consumption between 1949 and



2003, which in turn leads to an increase in the share of income that is spent on gasoline. Consumers tend to be more price sensitive with respect to consumer goods that take up a larger share of income.

**Table 5.5:** Estimation results from a meta-regression analysis based on a system of equations

Dependent variable:	$\epsilon_G$	$\epsilon_{FE}$	$\epsilon_{KM/C}$	$\epsilon_C$
	(1)	(2)	(3)	(4)
Constant	-0.107 (0.067)			
UCA	0.148 ** (0.038)	0.039 (0.436)	-0.016 (0.441)	0.204 ** (0.069)
Data Year / 100	-0.021 ** (0.005)	0.006 (0.022)	-0.013 (0.023)	-0.001 (0.005)
Oil_crisis	0.114 (0.078)	0.123 (0.207)	0.419 (0.272)	-0.181 (0.191)
Cross-section	-0.226 ** (0.724)	-0.057 (0.440)	-0.628 (0.481)	0.344 * (0.173)
Pooled	0.054 (0.039)	-0.155 (0.179)	-0.257 (0.184)	0.156 ** (0.052)
Long-run	-0.366 ** (0.047)	0.130 (0.109)	-0.327 * (0.130)	0.091 (0.083)
Dynamic	0.197 ** (0.041)	-0.295 ** (0.100)	-0.168 (0.104)	0.069 (0.057)
Non-linear	-0.182 (0.108)	0.105 (0.201)	-0.024 (0.765)	-0.053 (0.731)
Included variables	0.004 (0.006)	-0.001 (0.008)	0.027 * (0.012)	-0.024 * (0.010)
N		311		
Degrees of freedom		251		
Mean squared residual		0.030		

\* significant at the 5% level

\*\* significant at the 1% level

The results in the other columns enable us to draw conclusions as to how the results in Column 1 can be further interpreted in terms of the impact of study characteristics on price elasticities of fuel efficiency, car use and car ownership. For example, the result that price sensitivity with respect to gasoline demand is lower in UCA studies is mainly caused by significantly lower price sensitivity with respect to car ownership. With respect to fuel efficiency and car use, no significant difference is found between UCA studies and other studies. In the case of an increase in gasoline price, the interpretation is that UCA consumers respond to the increase in price by improving fuel efficiency and decreasing the use of existing cars to the same degree as their counterparts in other parts of the world. However, the

UCA consumers are less willing to decrease their number of cars. Because of this, changes in total gasoline demand caused by price changes are lower than in the rest of the world.

Looking at the coefficients of the trend variable, we see that there are negative time trends in the elasticities of total gasoline demand, fuel efficiency, car use and car ownership. However, only the result with respect to total gasoline demand is significant. The increased price sensitivity over time might be explained from the fact that the increase in gasoline consumption per capita has increased the share of the budget spent on gasoline. The effect of a change in the price of gasoline on the spending capacity has therefore increased over time. The use of data from the period following the oil embargo does not result in significantly different price elasticities.

Studies that use cross-section datasets show lower price sensitivity with respect to car ownership. The significant positive impact of the use of pooled data on the price elasticity of demand for gasoline is mainly caused by the significant positive impact on car ownership. All of these results indicate lower price sensitivity. As discussed before, the reason might be that time series estimates without extensive lag structures typically yield short term elasticity estimates while the estimation results from a cross-sectional study are more likely to be indicative of the long-run effect.

We find that in the long-run price sensitivity with respect to gasoline demand is significantly higher than in the short-run. The results show that this is mainly caused by higher price sensitivity of *car use* in the long run. We find no significant elasticities with respect to fuel efficiency and car ownership. This is unexpected, because the longer time span allows consumers more room for response and behavioral adjustment to the price change. This holds especially for behavioral adjustment related to car ownership and fuel efficiency.

From Column 2 we see that the use of a dynamic model significantly decreases the price sensitivity with respect to fuel efficiency. Because of the positive correlation between price and lagged fuel efficiency and the positive correlation between fuel efficiency and lagged fuel efficiency, this result is in agreement with the omitted variable formula (Greene, 2003, p.148). In a similar fashion, the positive coefficient of the dummy for dynamic models in Column 3 can be explained with the omitted variable formula. The use of a non-linear demand model does not have any significant effect on price sensitivity with respect to gasoline demand, fuel efficiency, car use or car ownership. Apparently, a loglinear model specification is sufficient to estimate the elasticity coefficients correctly.

## 5.6 Summary and Conclusion

The aim of this chapter is (i) to estimate the mean value for the price elasticity of gasoline demand and to decompose this into mean values for the elasticities of fuel efficiency, car use and car ownership with respect to gasoline price and (ii) to determine the impact of study characteristics and other conditional variables on the estimated elasticity values. In order to do so we collect estimates of six different types of price elasticities related to gasoline demand. Subsequently we investigate the relationship between the elasticity estimates and conditional variables in a meta-analytical way. In doing so, we utilize the theoretical relationships among the different types of price elasticities in order to enhance our results and improve their explanatory power. In Section 5.2, we describe the various channels through which a change in the gasoline price affects gasoline demand. We establish the theoretical relationship between the components of gasoline demand in relation to the concept of price elasticities. We discuss the linear relationships among the price elasticities of demand, fuel efficiency, car use, car ownership, traffic volume and gasoline consumption per car. In Section 5.3 we introduce a conceptual framework that enables one to combine observations on different effect sizes that are linearly related to each other.

In Section 5.4, we present some descriptive statistics of our dataset and we combined the estimates of the price elasticities in order to derive weighted mean elasticity values. We find a mean price elasticity of demand for gasoline of -0.53, which indicates that consumers are not very price sensitive to price changes. Furthermore, the results show that the response in demand for gasoline to a change in price is mainly caused by responses in fuel efficiency (-0.22) and car ownership (-0.22) and to a lesser degree to a response in car use (-0.10). Our results suggest that consumers are highly dependent on their car for transport and are not very willing to decrease the number of cars they own and even less willing to decrease the number of kilometers they drive with these cars. This suggests that fuel price charging alone is not a very effective instrument to reduce the external costs of individual road transport. Further results show that the response in demand to a change in the fuel price is mainly caused by responses in fuel efficiency and car ownership; price charges have a larger impact on fuel efficiency and the number of cars than on the number of kilometers per car. As flat fuel taxes alone do not seem to be very effective, they might better be integrated in a smart charging system that incorporates registration taxes and fuel efficiency tax-cuts or subsidization in order to stimulate the use of energy efficient technology. In this context, one could think of

innovative tax arrangements such as the "feebate" concept, based on taxes for vehicles with high fuel consumption along with rebates for vehicles with low fuel consumption, such as being introduced in some countries (Canada, Austria).

In Section 5.5, we investigate the relationship between the elasticity estimates and a set of moderator variables in a multivariate way. First, we conduct a conventional meta-regression analysis to explore the impact of study characteristics on the price elasticity of gasoline demand. The main results of the conventional meta-analysis – discussed in more detail in Section 5.5 - are as follows. We find no temporal variation in price elasticities of total gasoline demand. Furthermore, cross-section studies are found to report higher absolute elasticity estimates than time-series studies. The results confirm that long-run price elasticities are higher in absolute value than short-run estimates. The fact that gasoline demand becomes more price-sensitive in the long run needs to be acknowledged in the design of long run policy instruments. With respect to functional form, the use of a dynamic specification decreases the price sensitivity.

Subsequently we estimate a meta-regression model based on a system of equations in order to obtain more precise results and to explain the impact of study characteristics in a more precise and detailed way. The coefficients related to the impact of study characteristics on price elasticity of gasoline demand are slightly different, but the sign pattern is identical. The significance patterns are the same, except for the time trend variable, which enters significantly negative and the dummy for pooled datasets, which are significantly positive under SMR-estimation. In the design of long run policy instruments, this increase in price sensitivity over the years should be taken into account. Furthermore, the system of equations approach enables us to uncover some significant regional and temporal effects. We find that the lower price sensitivity with respect to total gasoline demand in the US, Canada and Australia is primarily due to lower price sensitivity with respect to car ownership. This result points to the high dependence of consumers on automobile transport and indicates that pricing policy could be more effective if fuel taxation is applied in combination with other types of charges such as registration fees or fixed charges on vehicle purchase. The increase in price sensitivity in the long run is mainly explained by an increase in the price sensitivity with respect to car use, which is something that policy makers should take into account. Functional form and model specification all have a significant impact on at least one of the investigated price elasticities.

The results demonstrate that the SMR-approach has several advantages. First, it allows for an integration of information on different types of elasticity estimates. As such, more observations can be used and combined than in the conventional approach. In the case of our dataset this leads to more precise results (i.e., lower standard errors). This is particularly useful in cases where we have a limited number of observations of a certain effect size. In these cases, the integration of information on related effect sizes might lead to rather different but significantly more precise results with respect to that particular effect size.

Furthermore, the SMR-approach enables the analyst to decompose the mean value of the elasticity of gasoline demand into mean values for the elasticities of fuel efficiency, car use and car ownership, and to identify impact effects that could not have been identified by using conventional meta-regression techniques. These findings help the analyst to investigate and interpret the relationship between gasoline demand and gasoline price in a more detailed and precise manner.

The SMR-model appears to be a useful estimation technique for meta-analyses in which the effect size of interest can be decomposed into partial effects that are of interest to the analyst. It appears to be a particularly useful approach if the effect size is an elasticity. First, the decomposition of an elasticity into a set of partial elasticities is usually based on a linear relationship.<sup>31</sup> This is convenient as it prevents the need for complicated non-linear estimation techniques. Second, an elasticity represents a causal relationship between a *variable of cause* and a *variable of effect*. If it is decomposed into a set of partial elasticities, these represent causal relationships between the variable of cause and a set of partial variables of effect that are functionally related to the variable of effect. Hence, the partial effects usually have a clear interpretation.

There are some restrictions to the application of the SMR model. First, observations for each of the partial effects are required. If observations are missing for one partial effect size, the estimation problem degenerates to a set of unrelated meta-regression models. In this case, the results with respect to the observed effect sizes are estimated separately. The results with respect to the unobserved effect size are calculated ex-post based on the relationship between the partial effects. If observations are missing for two or more partial effect sizes, the results with respect to unobserved partial effects cannot be estimated nor calculated. Second, the total number of observations should exceed the number of moderator variables multiplied

---

<sup>31</sup> This is because the decomposition of an elasticity is usually based on the decomposition of the variable of effect as the *product* of a set of partial variables. This is for example the case in the study on gasoline demand (see Equation (5.2)).

by the number of partial effects. A limited number of available observations for any individual effect size is not problematic, as long as the total number of observations is sufficiently large. If these conditions are satisfied, the SMR method is an appealing approach, which offers several advantages to conventional meta-analytical techniques.

## **Appendix 5: List of primary studies on price elasticities of gasoline demand**

- |                                  |                           |
|----------------------------------|---------------------------|
| Abdel-Khalek (1988)              | Greene (1990)             |
| Archibald and Gillingham (1980)  | Greene and Chen (1983)    |
| Archibald and Gillingham (1981a) | Houthakker et al. (1974)  |
| Archibald and Gillingham (1981b) | Kennedy (1974)            |
| Baltagi and Griffin (1983)       | Kraft and Rodekohr (1978) |
| Baltagi and Griffin (1997)       | Kwast (1980)              |
| Banaszak S et al.                | Lin et al. (1985)         |
| Bentzen (1994)                   | McRae (1994)              |
| Berndt and Botero (1985)         | Mehta et al. (1978)       |
| Berzeg (1982)                    | Mount and Williams (1981) |
| Blair et al. (1984)              | Ramanathan (1999)         |
| Dahl (1978)                      | Ramsey et al. (1975)      |
| Dahl (1979)                      | Reza and Spiro (1979)     |
| Dahl (1982)                      | Romilly et al. (1998)     |
| Donnelly (1982)                  | Samimi (1995)             |
| Drollas (1984)                   | Sterner (1991)            |
| Eltony (1993)                    | Tishler (1980)            |
| Eltony and Al-Mutairi (1995)     | Tishler (1983)            |
| Gallini (1983)                   | Uri and Hassanein (1985)  |
| Gately (1990)                    | Wheaton (1982)            |
| Gately (1992a)                   | Wirl (1991)               |
| Greene (1982)                    |                           |

## Chapter 6

# A Meta-Analysis of Technical Efficiency in Urban Public Transport<sup>32</sup>

## 6.1 Introduction

Urban transport, including bus, ferries, trams, light rail and metros, form a large share of the transportation network in an economy. Although travel patterns in most developed countries are increasingly dependent on the car (see Banister, 2000), causing a declining trend in transport demand in most industrial economies, urban transport remains an important transport mode. Urban transport services are provided by public, private or mixed companies in a highly regulated environment. Moreover, important constituents of the transportation infrastructure are essentially (semi-) public goods. There are sound economic reasons for a significant degree of state intervention in this field, based mainly on the recognition of a variety of market failures (see for example Kerstens, 1996). In the past two decades, however, serious concerns about possible regulatory failures have led to a reassessment of the role of the state in the organization of the sector (see Glaister et al., 1990).

In view of these concerns, it is of great interest to investigate whether urban transport operators work in a technically efficient way (i.e., reach economic targets such as cost minimization or output maximization conditional on output or input constraints). Solid technical efficiency measurement can provide a significant contribution to the discussion on the relative merits of private versus public provision of transportation services. From the early 1980s onward, various frontier estimation techniques have been developed to determine best practice behavior in an industry. Frontier methods allow for distinguishing between efficient and inefficient production and the estimation of the degree of (in)efficiency. In the transportation literature, frontier methods have been used in efficiency studies on almost all transport modes. A comprehensive survey of frontier methods and empirical results for urban

---

<sup>32</sup> This chapter is based on Brons et al. (2005).



public transport has recently been published (De Borger et al., 2002). While this survey provides a good overview of the literature on public transport performance, more extensive insights can be obtained by using quantitative-statistical research techniques.

The aim of this study is to present a statistical overview of the literature on public transport efficiency, and to give a statistical explanation for the variation in technical efficiency findings reported in the literature. Furthermore, we investigate if the results are indicative of regulatory failures and the benefits of deregulation of urban public transport.

In the next section, the concepts of technical efficiency and efficiency frontiers are introduced. In Section 6.3, we discuss the different frontier specification techniques that are used in the literature. Section 6.4 will discuss the feasibility of comparing technical efficiency studies in a meta-analytical set-up and will address some of the assumptions underlying this study. Section 6.5 consists of a statistical exploration of the methods and results that are found in the literature. In Section 6.6, we apply meta-regression analyses in order to identify determinants that may help explain the variation in efficiency results that are reported in the literature. Section 6.7 concludes with a brief summary and conclusions

## 6.2 The Concept of Technical Efficiency

Economists have traditionally distinguished between two sorts of efficiency: technical efficiency and allocative efficiency (Viton, 1986)<sup>33</sup>. In this study, we focus mainly on technical efficiency (TE). Technical efficiency relates to the divergence between actual production and production on the boundary of the feasible production set. This set summarizes all technological possibilities of transforming inputs into outputs that are available to the organization. A production process is technically efficient if production occurs on the boundary of the feasible production set and inefficient if production occurs within the interior of the set.<sup>34</sup>

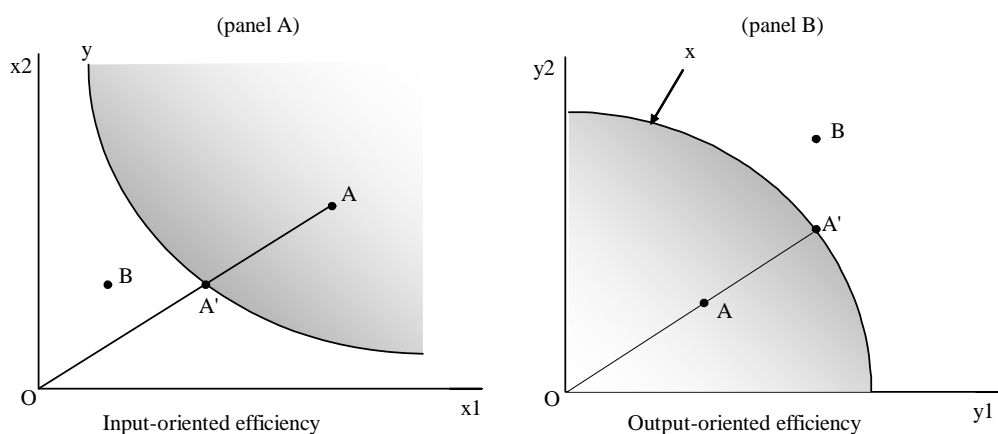
Technical inefficiency can be investigated from two perspectives (Viton, 1997). Input-oriented technical efficiency refers to the minimization of the amount of input commodities in order to produce a given amount of output while output-oriented technical efficiency refers to

---

<sup>33</sup> Other efficiency concepts frequently encountered in the literature include scale efficiency, which relates to a divergence between actual and ideal production size, and structural inefficiency, which relates to possible congestion of production regions (see De Borger et al. 2002).

<sup>34</sup> Allocative efficiency involves selecting that mix of inputs which produce a given quantity of output at minimum cost, given the input prices. Allocative and technical efficiency combine to provide an overall economic efficiency measure (Coelli et al. 1998).

the maximization of the output for a given amount of input commodities. Input-oriented technical inefficiency is illustrated in Figure 6.1 (panel A) for the case of two input commodities and one output commodity. In panel A, the shaded area represents the set of feasible input bundles  $(x_1, x_2)$ . The curve itself represents the set of efficient input bundles with respect to a given level of output. This set is called the efficiency or production frontier. Only locations on the curve are technically efficient. The input bundle represented by location A is technically inefficient because it uses more inputs than the input bundle represented by location A'. The degree of technical efficiency can be measured as the scalar  $(OA')/(OA)$ . This results in a ratio between zero and one. By construction, location B is not a feasible input vector with respect to the given level of output. In the literature, the calculation of input-oriented efficiency is sometimes based on the costs of the bundle of input commodities instead of the amount of input commodities. In this case, the efficiency frontier is called a cost frontier.<sup>35</sup>



**Figure 6.1:** Visual representation of input-oriented technical efficiency (Panel A) and output-oriented technical efficiency (Panel B)

Output-oriented technical inefficiency is illustrated in panel B for the case with two output commodities and one input commodity. The shaded area represents the set of feasible output bundles  $(y_1, y_2)$ . The curve itself represents the set of efficient output bundles (the efficiency frontier). The output bundle represented by location A is technically inefficient because the level of output is lower than in the output bundle represented by location A'. The degree of inefficiency can be expressed by the ratio  $(OA)/(OA')$ . By construction, output bundle B is not

<sup>35</sup> Note that in this case, the efficiency concept is based on a combination of technical efficiency and allocative efficiency.

feasible given the input level. Since the degree of technical efficiency can only be measured in relation to ‘best practice’, the efficiency frontier must first be constructed or estimated. This is the subject matter of the next section.

### 6.3 The Construction of Efficiency Frontiers

Efficiency frontiers can be constructed in an engineering context based on empirical knowledge of technical production operations (parametric specification). This can be done using either a deterministic or a stochastic approach. Furthermore, they can be estimated by observing production operations actually accomplished (non-parametric specification).

The deterministic parametric frontier method assumes a particular functional form for the efficiency boundary. A procedure that is called corrected OLS (COLS)<sup>36</sup> is used to measure technical inefficiency. First, an average practice frontier is estimated, using OLS. This frontier is corrected by shifting the intercept until all residuals except one become negative. The remaining non-negative residual should be equal to zero. Technical output efficiency is given by the ratio of the observation output value to the fitted frontier output value. Alternatively, one can directly estimate a frontier that envelops all observations by specifying an error term with a distribution that is truncated at zero, so that there is again at least one observation that coincides with the frontier.

The stochastic parametric frontier method is similar to the deterministic method, but allows for measurement error in the frontier. The error term consists a technical inefficiency component (deviation from the frontier) and a random error term with zero mean (measurement error of the frontier). Table 6.1 shows a categorization of frontier methods.

**Table 6.1:** A taxonomy of efficiency frontier methods used in public transport literature

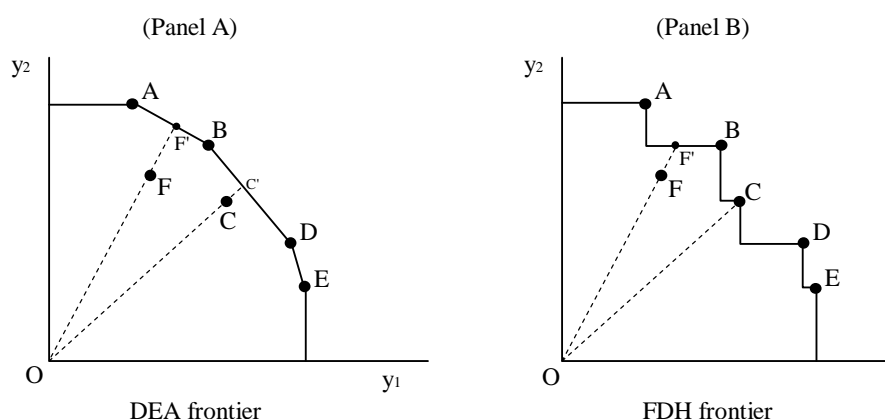
	Deterministic specification	Stochastic specification
<b>Parametric technology</b>	Frontiers based on (corrected) OLS or ML models.	Frontiers with explicit distributional assumptions for TE values
<b>Non-parametric technology</b>	Frontiers based on Free Disposal Hull (FDH) and Data Envelopment Analysis (DEA) techniques.	Resampling, chance constrained programming.

Source: Kerstens (1996)

<sup>36</sup> This method is sometimes referred to as displaced OLS.

Deterministic non-parametric methods are not based on a parametric production function, but on a piece-wise linear frontier that is directly constructed from the observations themselves by applying mathematical programming techniques. Two main approaches can be distinguished: (i) the Data Envelopment Analysis (DEA) method and (ii) the Free Disposal Hull (FDH) method (see Kerstens, 1996).

DEA modeling was initiated by Charnes et al. (1978). The DEA method provides relative measures of efficiency and is increasingly used in evaluating the performance of public service industries (for an overview we refer to Ganley and Cubbin, 1992). The efficiency measures are distances to an empirical production frontier and the values are calculated based on standard Pareto efficiency. No assumption has to be made about the production frontier's functional form, since the frontier is the observed best practice of the raw data set available. The frontier is constructed based on the assumption that any linear combination of observation units is feasible and on the assumption of strong input and output disposability.



**Figure 6.2:** Efficiency frontiers based on DEA assumptions (Panel A) and FDH assumptions (Panel B)

Strong input disposability means that a feasible output level remains feasible after increasing any input levels. Strong output disposability means that it is always possible to reduce the output level without changing input levels. The variable returns to scale DEA model is illustrated for output maximization in Figure 6.2 in panel A. Points A to E are observed output vectors. The model assumes that an observed output vector can be smaller than the linear combination of observations D and E. All observations to the south-west of the line segment D-E are therefore feasible. This explains the line originating in observation E and extending

parallel to the horizontal axis<sup>37</sup>. Using a similar reasoning for all line segments and allowing for all linear combinations yields the set of possible output combinations bounded by the production frontier that consists of the line segments A-B-D-E. Locations C and F are not efficient according to DEA assumptions. Hence, they are not on the frontier. The degree of technical inefficiency for locations C and F is measured by the fraction  $(OC)/(OC')$  and  $(OF)/(OF')$ , respectively. Strong input disposability can be illustrated analogously.

In FDH models (initiated by Deprins et al., 1984) the efficiency frontier is constructed based on the assumptions of strong input disposability and output disposability but without allowing for linear combinations of observational units. Based on these assumptions the FDH frontier typically has a staircase form, as illustrated in panel B. Note that the set of observed locations is identical as that in panel A. The set of feasible outputs is bounded by the production frontier that consists of the line segments A-B-C-D-E. Thus, location C, which was inefficient in the DEA model, is efficient in the FDH model. Location F remains inefficient under FDH assumptions. The degree of inefficiency of location F is measured by the ratio  $(OF)/(OF')$ . Since FDH models relax the assumptions of DEA models the observational units are on average located closer to the frontier<sup>38</sup>. Therefore, all other things equal, technical efficiency ratios are expected to be lower for studies based on DEA modeling than for studies based on FDH modeling<sup>39</sup>. The construction of the production frontier based on strong input disposability can be illustrated analogously.

Stochastic non-parametric frontiers have, to our knowledge, not yet been used in the literature on public transport efficiency performance.

## 6.4 Study Comparability Issues

The primary studies on which this meta-analysis is based, report mean TE values; i.e., the mean of a sample of observed TE for individual firms (cross-section data), time periods (time series data) or both. Note that one underlying study can yield multiple observations<sup>40</sup>. An important issue that needs to be acknowledged is that technical efficiency is a *relative*

<sup>37</sup> Note that, by definition, any point on the efficiency frontier between E and the  $y_1$ -axis is efficient, although the production bundle is dominated by the bundle in point E for all positive prices of  $y_1$ . Such differences between optimal and non-optimal points on the frontier are referred to as 'slacks'. The undesirable properties of slacks can be handled by adding an additional coefficient to the linear programming problem, which prevents the frontier from running parallel to the axis.

<sup>38</sup> Note that the set of observation points is the same in panel A and panel B. Point C, which is located on the efficiency frontier under FDH assumptions, is not efficient under DEA assumptions.

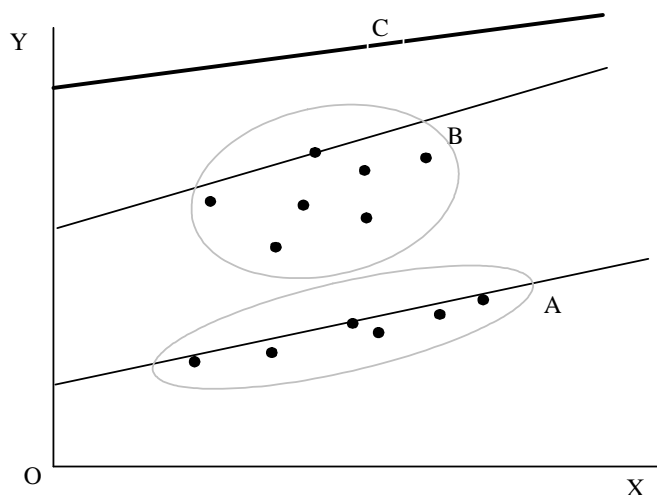
<sup>39</sup> This can easily be seen by comparing the efficiency of point F under DEA and under FDH assumptions.

<sup>40</sup> A list of primary studies used in this study is given in Appendix 6.

measure. The reference point for measuring TE is a case-specific efficiency frontier, which is determined based on the actual observation points within the sample. Therefore, the efficiency frontier itself may not be fully efficient.

Under the theoretical assumption that there exists an efficiency frontier with universal validity, it is not (in)efficiency values that we are comparing but rather sample heterogeneity. The issue is illustrated in Figure 6.3 for the one-input-one-output case under parametric assumptions. The homogeneous sample A will yield a higher mean TE value than the heterogeneous sample B. Yet, if we assume that there is an overall efficiency frontier, represented by C, it is obvious that the real TE is higher for sample B than for sample A<sup>41</sup>.

Alternatively, under the assumption that the study-specific efficiency frontiers indeed represent the maximum attainable efficiency relations given the study-specific conditions we can interpret the comparison of mean TE values as the comparison of actual efficiency scores<sup>42</sup>.



**Figure 6.3:** Illustrative case of the problem of comparing mean TE values from different studies

Obviously, neither of the two assumptions would hold in practice. Firms that are ‘efficient’ within a specific sample might still be able to do better, given the environmental conditions<sup>43</sup>. On the other hand, the observed TE values of the sample *do* provide some indication of the maximum attainable efficiency under the case specific conditions. Comparison of reported TE

<sup>41</sup> In the figure the frontiers are located as they are for illustrative purposes. Note that in practice we do not directly observe the difference between intercepts (although under certain conditions they can be derived indirectly).

<sup>42</sup> Comparison is also possible under the assumption that the *actual* (non-observed) TE of ‘efficient’ firms is similar among studies.

<sup>43</sup> This includes conditions related to the institutional, cultural, geographic and economic environment.

values therefore implies the comparison of values that consist of a heterogeneity component and a technical efficiency component.

There are several ways to deal with this problem. One solution is to assume revealed optimal efficiency (i.e., the observed frontier is the actual frontier). Obviously, this is a strong assumption.

A weaker assumption, based on the notion that the observed efficiency values are indicative of the highest possible efficiency given the circumstances (i.e. the observed frontier is related to the actual frontier) is that there is a fixed linear relation between the observed and the actual frontier, which is identical for all observations. In other words, the possibilities for improving efficiency are similar for all best practice firms within a sample. Clearly, the strong assumption of revealed optimal efficiency is a special case of the weaker assumption of a fixed linear relation between revealed and optimal efficiency, viz. the case where constant and coefficients are zero.

The previous discussion suggests that, although comparing average TE values between studies may lead to biased results, comparing average TE values that come from the same study would be less problematic<sup>44</sup>. Therefore, another way to deal with the problem of study comparability is to specify the meta-regression equation such that the variation in the TE estimations is only explained by the within study variation. This can be done by the using dummy variables for different studies. Obviously, the efficacy and feasibility of such a solution depends on the average number of observations per study.

A third way to circumvent the comparability problem involves a re-interpretation of the dependent variable. While the actual technical efficiency frontier is not observed, the average measured TE value does provide an indication of the relative variation in technical efficiency values, and thus of the possibilities to improve technical efficiency<sup>45</sup>. If the dependent variable is interpreted as such, comparison between studies in a meta-analytical format is indeed valid. The requirement that the individual firm with the best efficiency within a study have a TE value of one is not a necessary condition for interpreting individual TE values as an indicator of relative efficiency within a study.

Finally, assuming that there indeed is some universal efficiency frontier one could transform the study specific efficiency frontiers in such a way that they become directly

<sup>44</sup> Note that this only holds if the TE values from one study are all based on the same set of observations or are all based on different subsets of the same set of observations.

<sup>45</sup> Consider the following measure of variation:  $\sum_n |1 - x_n|$ . Because  $x$  in our case is a measure of efficiency, it is positive and smaller than or equal to 1. Dividing this measure by  $n$  then yields  $1 - \bar{x}_n$ , which yields the same information as  $\bar{x}_n$ , the average TE-value.

comparable. Theoretically, this could be possible for parametric frontier studies if one has sufficient knowledge about the intercept and slope of the input-output curve (in the absence of such data information on input and output levels of all the firms in the original studies is needed). In practice, such data is missing for a great number of studies. Moreover, such transformation calculations become increasingly complicated when the number of input and output goods increases and even more so when assumptions about the functional form of the production technology differ among studies. For non-parametric studies, such transformations are usually not possible in the first place because the frontier is not expressed in terms of a functional form.

Since the database underlying this study contains multiple studies with only one observation, the remedy of focusing on the within-study variation only is not feasible in practice. Too much statistical information would be discarded and the econometric model would collapse due to near-singularity of the regression matrix. In addition, the transformation of the dependent variable values is not feasible in practice, mainly due to a lack of information. The assumption of revealed optimal efficiency (or a fixed linear relationship) may be too strong. The reinterpretation of the dependent variable has similar implications for the further analysis: no additional *statistical* procedures are needed to remedy the comparability problem. Therefore, in the remainder of the study we interpret the mean technical efficiency value as a measure of relative technical efficiency. When discussing the results of the empirical analysis in this study, we pay attention to the close link between the mean relative efficiency value and the variation in individual efficiency values from the underlying study.

## 6.5 Quantitative Overview of the Literature on Public Transport Efficiency

In order to discuss the literature on public transport efficiency we use a sample of parametric and non-parametric case studies. Most of these studies have been collected and used for a survey by De Borger et al. (2002). For a qualitative description and discussion of those studies, we refer to this study. In Table 6.2, we present a quantitative summary of the literature. The average technical efficiency ratio of the total set of studies is 0.825. This ratio is slightly higher for parametric studies (0.847) than for non-parametric studies (0.814).

Furthermore, we see that 70 per cent of the parametric studies use panel data, either balanced or unbalanced. Among the non-parametric studies, this is about 50 percent, all of



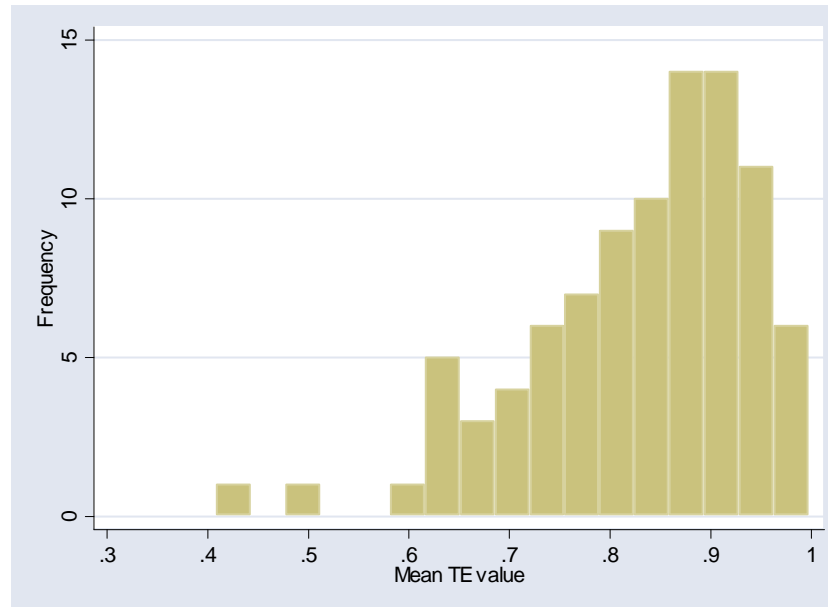
them being balanced panel data sets. Apparently, among non-parametric studies cross-section databases seem to be preferred. About 60 per cent of the studies use European data. In about 25 per cent of the studies, U.S. data are used. Parametric studies use relatively more data from Asian countries while non-parametric studies use relatively more U.S. data. Two thirds of the parametric studies use stochastic frontier methods. Only 33 per cent use deterministic specifications. In two thirds of the parametric studies, cost frontiers are used to derive technical efficiency. Production frontiers are used in the remaining studies. Most of the non-parametric studies use DEA techniques in order to construct an efficiency frontier (89 percent). FDH assumptions are used in the remaining 11 per cent of these studies.

**Table 6.2:** Statistical overview of characteristics of the case studies used in the meta-analysis

	Parametric studies	Non-parametric studies	Total sample
<b><u>General</u></b>			
Number of studies	15	18	29
Number of observations	30	63	93
Average TE	0.847	0.814	0.825
<b><u>Database type</u></b>			
Cross-section	16.7%	36.5%	30.1%
Time series	13.3%	14.3%	14.0%
Panel	60.0%	49.2%	52.7%
Unbalanced panel	10.0%	0.0%	3.2%
<b><u>Region</u></b>			
Europe	56.7%	60.3%	59.1%
USA	10.0%	33.3%	25.8%
Asia	33.3%	6.3%	15.1%
<b><u>Parametric specifications</u></b>			
Stochastic	66.7%		
Cost frontier	66.7%		
<b><u>Non-parametric specification</u></b>			
DEA		88.9%	
FDH		11.0%	
<b><u>Output indicator related to</u></b>			
Passengers	50.0%	22.2%	31.2%
Seats	40.0%	9.5%	19.4%
Vehicles	16.7%	69.8%	52.7%
Revenues	3.3%	22.2%	16.1%

There is a large variability in the use of output measures, suggesting that there is no generally accepted set of appropriate variables in the urban transport sector. Most studies use a

combination of different output measures. The majority of the measures used are related to numbers of passengers, seats and/or vehicles<sup>46</sup>. Table 6.2 shows that in parametric studies passenger and vehicle related output measures are more often used than in non-parametric studies. The latter focus more frequently on vehicle related output measures. In non-parametric studies, combinations of different output measures are used more often than in parametric studies.



**Figure 6.4:** The distribution of mean TE values

Figure 6.4 shows the distribution of technical efficiency ratios of the complete sample. The overall mean technical efficiency value is 0.825. The distribution of the ratios is skewed to the right with a peak around 0.9. The shape of the distribution indicates that, on average, there are relatively few highly inefficient companies; in general, urban transport companies tend towards efficiency.

Table 6.3 shows the average technical efficiency ratios for different subsets of the set of public transport efficiency observations. The observations are categorized according to study characteristics such as database type, geographic region, deterministic versus stochastic frontier specification, cost versus production frontier and non-parametric model type. As the table shows, the use of different study characteristics does generally not lead to large

<sup>46</sup> For example, the dummy variable “passengers” is a dummy for studies that include in the econometric model an output measure related in one way or another to passengers (e.g, number of passengers per km (or mile), number of passengers per month etc.).

differences in efficiency ratios. Parametric studies that use stochastic frontier specifications find a slightly lower degree of efficiency on average compared with deterministic studies, which is somewhat surprising as deterministic studies attribute all errors to inefficiency. Furthermore, the set of parametric cost frontier studies shows a slightly lower efficiency ratio compared to the set of parametric studies that use production frontiers. This could be due to the fact that cost frontier inefficiencies include both technical and allocative effects. Observations based on time-series analysis and observations from studies based on U.S. data generally find higher relative efficiency values than their reference categories. This holds for both parametric and non-parametric sets of observations. Among the non-parametric studies, DEA studies report on average slightly lower relative efficiency levels than FDH studies. This can be explained from the fact that in DEA studies assumptions with respect to input-output relations are stronger than in FDH studies. This results in the construction of a frontier that is more efficient in terms of the input-output relationship. All other things being equal, observational units in DEA studies on average will be located farther away from the frontier than observational units in FDH studies.

In addition to the summarized relationships between technical efficiency and the various categorical variables as shown in Table 6.3, it might be interesting to investigate the univariate relationship between efficiency and some of the continuous variables that we use in our database by means of various scatter plots.

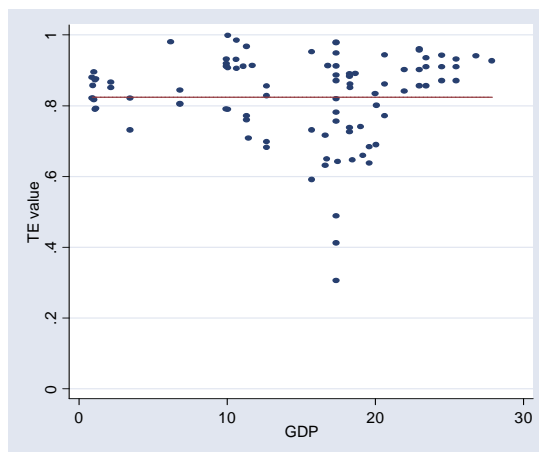
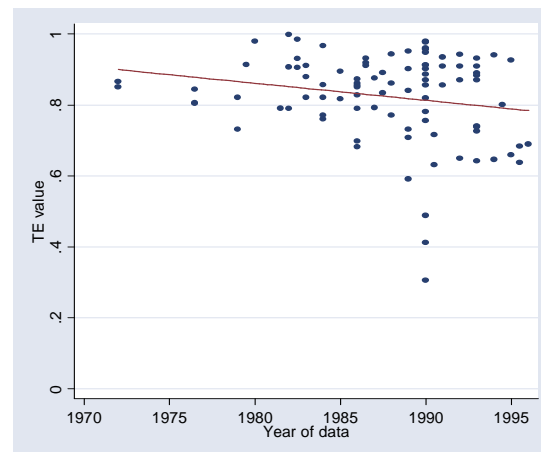
In Figure 6.5 the univariate relationship between GDP per capita and the technical efficiency ratio is shown. The correlation coefficient is approximately zero. Apparently, higher GDP and the associated increased car ownership, which in turn implies a lower demand for public transport does not affect efficiency ratios.

From the scatter plot in Figure 6.6, we see that the correlation between the year of data and technical efficiency is negative. As the previous scatter plot showed, this cannot be explained from the GDP growth over time and associated consequences for urban transport demand. Another reason for the negative sign of the latter coefficient might be that due to technological innovation and outsourcing the frontiers may have shifted to new efficiency standards. Thus, when the technical and institutional environments of urban transport become more heterogeneous, one may expect that variations in efficiency increase. Rigid public transport organizations may not have adjusted fully to these new standards.

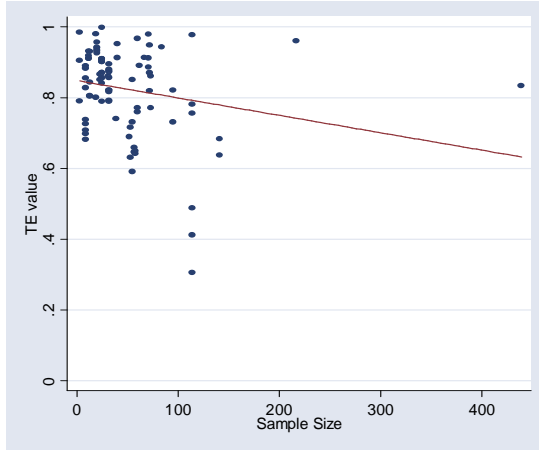
**Table 6.3:** The average TE values for various sub-sets of the sample of observations

	Parametric studies	Non-parametric studies	Complete sample
complete set of studies	0.847	0.814	0.825
cross-section studies	0.823	0.767	0.777
time series studies	0.919	0.895	0.903
panel data studies	0.847	0.825	0.833
unbalanced panel studies	0.791		0.791
European studies	0.842	0.826	0.792
USA studies	0.896	0.904	0.893
Asian studies	0.841	0.843	0.834
stochastic studies	0.836		
deterministic studies	0.870		
cost frontier studies	0.836		
production frontier studies	0.869		
DEA		0.809	
FDH		0.847	
Studies using passengers as output measure			0.802
Studies using vehicles as output measure			0.837
Studies using seats as output measure			0.821
Studies using revenues as output measure			0.844

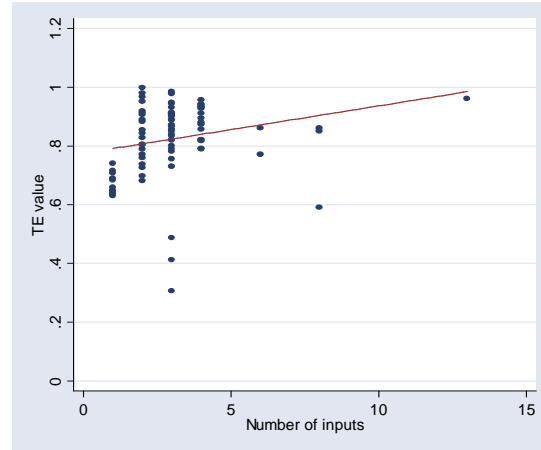
Due to the nature of the dependent variable (i.e., the interpretation as a distribution measure), it is interesting to investigate the relationship between the mean technical efficiency value and the sample size of the underlying study (see also Zhang and Bartels, 1998). Figure 6.7 shows the correlation between the two variables. The correlation is negative but not very strong at -0.22.

**Figure 6.5:** The correlation between technical efficiency and GDP per capita**Figure 6.6:** The correlation between technical efficiency and year of data

In Figure 6.8, the mean TE values are plotted against the number of inputs that are used in the primary studies to estimate technical efficiency. The correlation coefficient is positive which can be explained from the fact that the use of more inputs in the econometric model increases the explained variation in the systematic part of the cost or production function. Because of the way the mean TE value is constructed, an increased number of inputs may therefore lead to higher values.



**Figure 6.7:** The correlation between technical efficiency and sample size



**Figure 6.8:** The correlation between technical efficiency and number of inputs used in the original study

## 6.6 Results of Meta-Regression

One of the aims of this study is to provide a statistical explanation for the variation in technical efficiency findings reported in the literature. In order to identify moderator variables which cause such variation we carry out two meta-regression analyses. Compared to the exploratory analysis in the previous section the multivariate techniques in this section provide us with more possibilities to correct for various factors while assessing the partial effects of certain variables. For this analysis, we pooled the sets of parametric and non-parametric studies and used a dummy variable to correct for fixed effects between the two sets. Furthermore, we use the following variables to explain the variation in the mean TE values: dummy variables for time series studies and panel data studies<sup>47</sup>, a trend variable<sup>48</sup>, GDP per capita, the share of government expenditure as a percentage of GDP, a dummy variable for stochastic parametric studies, dummies for U.S. and Asian data, a dummy for studies that use

<sup>47</sup> This dummy refers to both balanced and unbalanced panel data

<sup>48</sup> The trend variable is calculated as the average year of the data on which the observations is based.

cost frontiers, a dummy for non-parametric DEA studies, a dummy for parametric studies, dummies for the type of output measures that are used in the estimation model, the sample size and the number of inputs<sup>49</sup>.

In the econometric specification, we use a transformation of the mean technical efficiency values as a dependent variable so that the estimated values fall within the range between zero and one<sup>50</sup>. We estimate the model both unweighted and weighted for sample size. For both estimations, we use the multilevel variance components estimator, discussed in Chapter 4, to correct for within-study dependence among observations from the same study<sup>51</sup>.

Table 6.4 shows the results of both estimations. The results in column 1 show the unweighted results. We see that the type of database that has been used in the original study affects the mean TE value. TE values from time series and panel data studies are significantly higher than those from cross-section studies. Apparently, increased technical efficiency, primarily due to technological change, has resulted in a variation in TE values that is large compared to the variation in TE values of a sample of firms at a specific moment in time.

The significant positive coefficient of the time trend variable indicates that technical efficiency has increased over time. The reason might be that, from the late 1970s onwards, many transport activities have been returned to private hands under the banner of privatization and deregulation. This indicates that the efficiency of public transport operators would benefit from a deregulated environment and thus that the concerns about possible regulatory failures are valid.

The coefficient for the government share variable is significantly negative. Under the assumption that the government share is related to the degree of government intervention (including subsidization) in urban public transport, this result also supports a deregulated environment with respect to urban public transport. The negative and significant coefficient of GDP indicates that welfare level and technical efficiency are negatively related to each other. This could be explained from the fact that government share rises with per capita income in

---

<sup>49</sup> By using dummy variables for parametric studies, stochastic studies and DEA studies, we basically follow the taxonomy of methodologies in Table 8.1. Note that these dummies are not directly related to each other in the sense that they aim to form a mutually exclusive set, so that multicollinearity may be a problem. We experimented with an alternative dummy structure that is based on the notion that ultimately there are four frontier methodologies that can be made directly comparable by including dummies for three of them and using the remaining one as a reference category. The results of such a dummy structure are essentially the same as the one we use here. We extended this experiment by further dividing the methodologies based on the distinction between cost- and production-based studies, resulting in six frontier methodologies and thus five dummies. The results were very similar to our findings (results are available on request).

<sup>50</sup> We use the transformation  $\ln[y/(1-y)]$  to estimate the following model:  $y = \exp(X'\beta)/[1 + \exp(X'\beta)] + \mu$

<sup>51</sup> The likelihood test based on the difference in deviance for the least squares model and the multilevel model is 9.22 for the unweighted model and 8.70 for the weighted model. This indicates that in both models significant within-study correlation is present. The within-study correlation coefficients are 0.69 and 0.74, respectively.

both cross-section and time series estimations (see for example Easterly and Rebelo, 1993 and Oxley, 1994), although we already included government share in the regression. Furthermore, if welfare level increases, transport operators focus increasingly more on transport characteristics such as service quality, comfort and speed instead of only on cost minimization and low prices. This may not lead to lower revenues due to higher ticket prices, but it might lead to lower technical efficiency if output is measured in number of seats, vehicles or passengers.

Having corrected for GDP level we see that US studies report higher TE values compared to European studies. Perhaps the reason might be that in a market like the US, where the degree of deregulation and privatization is higher and subsidies are lower, the relative importance of profit maximization versus accessibility as the industry's goal is higher. Since car ownership is higher in the U.S. and the infrastructure accommodates private transport to a larger degree, public transport may be less developed on low profit segments of the public transport network. Technical efficiency is significantly lower in Asia, which is mainly represented by Taiwan and India in our dataset. This may be caused by the large government share in Taiwan and India, even though we already corrected for government share.

The dummy variable for DEA studies is significantly negative<sup>52</sup>. As previously discussed, this makes sense because of the stronger frontier assumptions of DEA studies compared to FDH studies. There appears to be no significant difference between parametric and non-parametric studies. The dummies related to output measures enter significantly negative. The reason may be that the inclusion of revenues in the estimation model leads to a better goodness of fit; the reference category here contains primarily studies that use revenues as an output measure. The number of inputs variable enters positive and significant, which confirms our expectations about a positive relationship between goodness of fit and the mean TE value. Finally, the coefficient of the sample size variable enters insignificantly.

The second estimation uses the same set of explanatory variables but uses an estimation procedure to weight for sample size<sup>53</sup>. The results are in column 2. Compared to the non-weighted results, the coefficients of GDP, government share and Asia, become insignificant in the weighted model. This means that the support for deregulation of urban

---

<sup>52</sup> Although formally the reference category is non-DEA studies, we can interpret the result as if the reference category would be FDH studies. This is because we use a dummy to correct for parametric studies and both DEA and FDH studies are non-parametric studies.

<sup>53</sup> By using the sample size as a weight variable, more weight is given to more precise estimates. It is common practice in meta-analysis to weight for the inverse of the variance of an effect size. However, for a large number of observations these were not available and could not be calculated.

public transport decreases somewhat, although some of the variables that are indicative of the benefits of deregulation, i.e. the U.S. dummy and the time trend variable, remain significant in the weighted model. The dummy for observations that use seats as output measure becomes insignificant. The sample size enters significantly negatively in the weighted estimation. This result corresponds with the correlation between TE value and sample size shown in Figure 6.8.

**Table 6.4:** Estimation results of the meta-regression analysis

	(1)		(2)	
	Coefficient	Std. Error	Coefficient	Std. Error
(Constant)	-75.667	39.071	-9.548 **	4.618
Time series	1.719 **	0.318	3.757 **	1.150
Panel	1.204 **	0.282	2.729 **	0.670
Datayear	0.041 *	0.020	0.236 *	0.114
GDP	-0.113 **	0.035	-0.195	0.159
Gov_share	-0.065 *	0.025	-0.047	0.094
USA	1.934 **	0.595	2.897 **	1.370
Asia	-0.830 *	0.406	-0.332	1.561
Stochastic	-0.528	0.571	0.315	1.099
Cost frontier	-0.605	0.623	-1.901	1.051
DEA	-1.315 **	0.177	-1.509 **	0.201
Parametric	0.093	0.690	0.210	1.043
Passengers	-1.112 **	0.369	-1.970 **	0.434
Vehicles	-1.237 **	0.272	-1.876 **	0.185
Seats	-1.120 **	0.355	-0.808	0.703
Number of inputs	0.263 *	0.129	0.568 **	0.111
Sample size	-0.006	0.003	-0.011 *	0.002
Between study variance	0.77		1.19	
Within-study variance	0.35		0.43	
Within-study correlation	0.69		0.74	
N	93		93	
log-likelihood	-116.22		-127.07	

\* significant at the 5% level

\*\* significant at the 1% level

## 6.7 Summary and Concluding Remarks

The aim of this study is to present a statistical overview of the literature on public transport efficiency, and to give a statistical explanation for the variation in technical efficiency findings reported in the literature. Furthermore, we investigate if the results are indicative of regulatory failures and the benefits of deregulation of urban public transport.



In Section 6.2, we introduce the concepts of technical efficiency and efficiency frontiers. In Section 6.3, the frontier specification techniques that are used for technical efficiency analysis are discussed. In Section 6.4, we discuss the feasibility of comparing technical efficiency studies in a meta-analytical set-up and made some important assumptions underlying this study. Section 6.5 presents the results of a statistical exploration of the dataset that we constructed based on the literature. In Section 6.6, we carry out meta-regression analyses in order to identify determinants that affect the technical efficiency estimate. The results are presented in Table 6.4.

The results of the study show that technical efficiency is similar for parametric and non-parametric studies. Estimations based on time series or panel data result in a higher mean TE value than cross-section studies. As expected, DEA assumptions lead to lower efficiency values than FDH assumptions. The variables that are used to measure output also affect the TE value; the use of passenger-, seats-, and vehicle-related output indicators leads to lower efficiency estimations compared to studies based on revenue measures. The sample size of the original case study does not affect the TE value when using the unweighted model but has a significantly negative coefficient when using the weighted model. The number of inputs that are used in the estimation model does not have a significant impact on the TE value. Among parametric studies, there is no significant difference between stochastic and deterministic specifications. The choice between a cost frontier and a production approach also does not affect the TE value when using the unweighted estimation model.

More importantly, the unweighted results imply that a deregulated environment is beneficiary for the efficiency of public transport operators. The negative coefficients for GDP and government share and the positive coefficients for the U.S. dummy and the time trend variable directly or indirectly suggest a negative relationship between state regulation and technical efficiency. When using the weighted estimator, the GDP and government share variables lose significance. The U.S. dummy and the time trend variable, however, remain significant. Obviously, the concerns about possible regulatory failures, mentioned in the introduction to this chapter, are justified.

**Appendix 6: List of primary studies on technical efficiency of urban public transport**

Bhattacharyya et al. (1995)	Levaggi (1994)
Button and Costa (1999)	Lijesen (1998)
Chang and Kao (1992)	Loizides and Giahalis (1995)
Chu et al. (1992)	Matas and Raymond (1998)
Costa (1998)	Nolan (1996)
Costa and Markellos (1997)	Nolan et al. (2002)
Cowie (2002)	Obeng (1994)
Cowie and Asenova (1999)	Sakano and Obeng (1995)
De Jong and Cheung (1999)	Sakano et al. (1999)
Delhause et al. (1992)	Thiry and Tulkens (1992)
Fazioli et al. (1993)	Tulkens (1993)
Filipini (1992)	Viton (1986)
Gathon (1989)	Viton (1997)
Hanusch and Cantner (1991)	Wunsch (1994)
Kerstens (1996)	



# Chapter 7

## Rail Noise. Economic Valuation and Policy<sup>54</sup>

### 7.1 Introduction

Both economic growth and land use policy have caused a situation where noise pollution, from both surface and airborne traffic, forms an ever-increasing burden on the residential environment. Noise does not only generate a reduction of the sense of well-being of those affected but also causes depreciation of property values. As a result, noise disturbance has become one of the most important forms of environmental pollution in industrialized economies. Noise pollution is an economic externality; not all social costs related to noise are reflected in the price of good. As silence does not have a market price, it is necessary to deduce its price indirectly. Hence, establishing an appropriate compensation fee is a complicated matter.

In many countries, the use of public transport (in particular, mass transport systems) is favored to decrease the negative environmental effects of private transport (apart from the equity elements involved). In order to stimulate the use of public transport, governments tend to plan residential areas in the proximity of railroad terminals or rail infrastructure, while at the same time accessibility of these areas is increased by expanding the rail network. Due to this policy, rail noise disturbance has recently become an issue of increasing importance. State-of-the-art cost-benefit analyses should take into account the increasing costs that are associated with noise disturbance. To that purpose, a sound valuation technique and a greater insight on the phenomenon of rail noise at large are needed.

Rail noise is a complex phenomenon; many interrelated concepts such as emission values, immission values, noise disturbance and economic costs play an important role. The purpose of this study is to investigate the relationships between these concepts, and to identify opportunities for the government to use the information on these relationships to reduce the external costs of noise from rail transport. Furthermore, the trade-off between damage costs

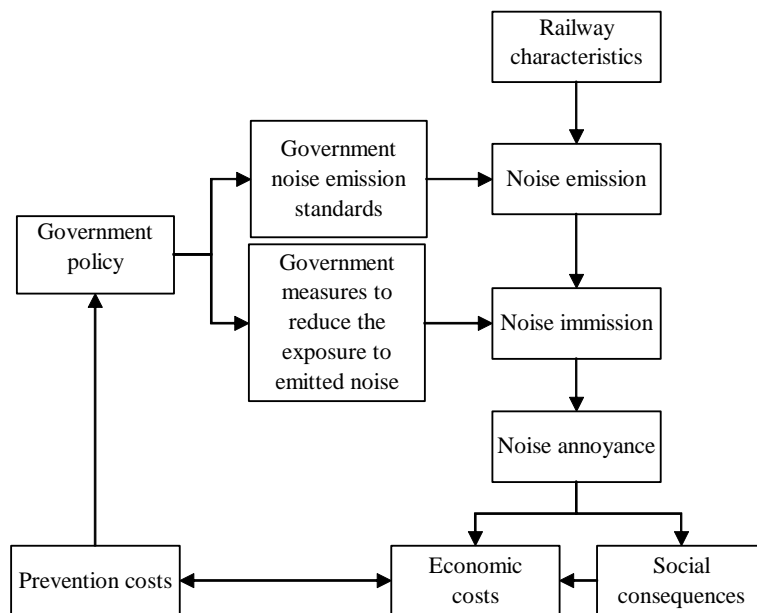
---

<sup>54</sup> This chapter is based on Brons et al. (2003).

and noise prevention by the government is discussed. This includes a literature survey on the valuation of rail noise pollution.

## 7.2 Rail Noise, External Costs and Policy: a Conceptual Model

Rail noise is an interdisciplinary problem; both economic systems and processes and environmental issues are involved. Economic commodities can only be converted into other economic commodities by means of a co-transformation of natural resources into (noise) emissions (see Heijungs, 2001). Figure 7.1 shows a conceptual model of economic and environmental interactions and causalities related to rail noise, in which government policy is a central factor. The model has a feedback loop that links the economic costs of rail noise with cost reduction policy. In the model, the generation of noise emissions is determined by railway traffic characteristics such as frequency and speed, and by limit values, which are determined by government policy.



**Figure 7.1:** A conceptual model of the causal relationship between rail noise, external costs of noise and policy

Noise emission values and immission values are not necessarily equal to each other<sup>55</sup>. The difference between these values is determined by the distance between the railroad track and the measurement point, meteorological factors, and the presence of objects, located between

<sup>55</sup> The emission value is the level of noise, measured in dB(A), at the noise source. The immission value is the level of noise, measured in dB(A), at the location of impact, such as a residential building or any other receiving property.

the railroad track and the measurement point, that interfere with the noise dispersion. Noise control barriers are an example of the latter category. The relationship between noise immission and noise disturbance is affected by, for example, activity patterns, attitude towards the railway and habituation to noise.

Noise disturbance has detrimental social and economic consequences. Social effects involve both psychological and physiological health problems. Economic effects are manifold and diverse but they are always costs. Social costs may lead to economic costs; it is obvious that school buildings, medical premises, residential areas and business premises exposed to noise affect the economy through the human capital stock. Rail noise also has a negative effect on property values. Moreover, noise limit values put restrictions on construction plans in the vicinity of the railroad track. There are monetary costs involved in the restrictions to such economic development or in efforts to comply with noise limit values. Sometimes, the feasibility of noise reduction measures is assessed by a cost-benefit analysis.

The individual components of the conceptual model presented in this section are discussed in more detail in sections 7.3, 7.4, 7.5 and 7.6.

### **7.3 Government Policy: Emission Standards**

Government policy on noise disturbance is primarily directed along two lines of measures, i.e. regulation and noise reduction policy. The first category includes regulation with respect to noise emission and immission standards, but also with respect to the methods that are used to measure noise. The second category includes policy measures that aim to reduce noise emission and immission, and incentives to private agents, such as railway operators and residential developers, to apply such measures. An example is the construction of noise control barriers. Noise reduction measures are discussed in section 7.5. In the present section, we describe government policy on noise emission and immission standards.

Noise legislation offers governments various instruments to reduce noise emission and immission values. These include restrictions with respect to noise emissions from rail vehicles, restrictions with respect to the temporal distribution of railway traffic and restrictive conditions with respect to the construction of the railway infrastructure. Sometimes a zone regulation system is used that is similar to the one used for highways. Such a zone regulation system establishes a zone along every railway line, whose width varies from 100 to 500

meters, depending on traffic density. Within such a zone, rail noise may not exceed certain limit values.

Different limit values apply to e.g. hospitals, schools and business premises (see Table 7.1 for the Netherlands). These limit values are relatively easy to impose when constructing new railway lines or buildings. In the case of existing urban areas and railway lines, additional measures related to vehicles and infrastructure are required. Note that in many countries the simultaneous development of urban areas and of railway networks in the nineteenth century has led to situations with high noise levels close to existing buildings.

**Table 7.1:** Limit values for rail noise in dB(A) for different building types in the Netherlands

Building type	day	evening	night	24 hours
Noise sensitive buildings (schools, hospitals)	55	50	45	55
Residential buildings	57	55	50	57
Office buildings	65	60	55	65

Source: De Jong (2001)

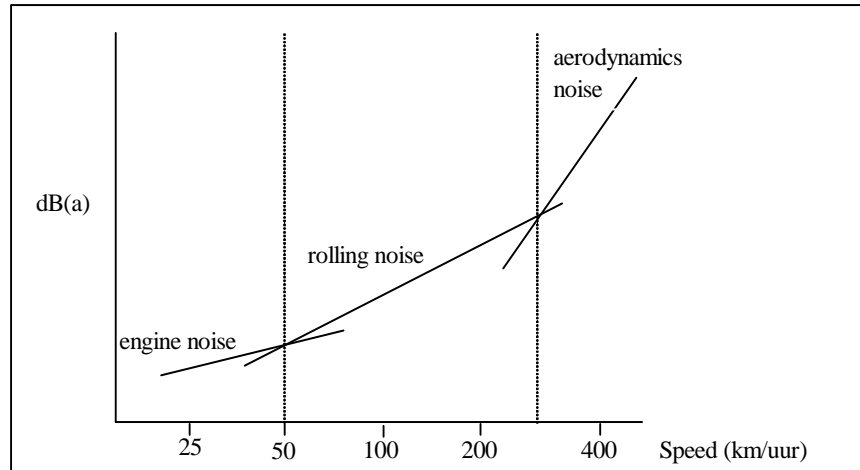
## 7.4 Noise Emission and Immission

There is a close, but complex, relationship between the emission and the immission level of noise. Government policy that aims to reduce noise disturbance can be based on the reduction of noise emission and/or noise immission levels.

### 7.4.1 Noise Emission Sources and Reduction Measures

Noise emission is initially determined by railway system characteristics, such as traffic density, frequency, speed, train type and rail-infra-structural characteristics. Specific noise emission sources from railway transport can be categorized into six categories: (i) rolling noise from vehicles on straight rails without discontinuities; (i) bumping noise from discontinuities on wheels or rails (such as intersections and junctions); (iii) noise from vehicles passing through a curve; (iv) noise generated by diesel engines; (v) aerodynamic noise caused by turbulence; (vi) other sources such as braking, railway maintenance, station noise or crossroads warning signs. Figure 7.2 shows that there is a positive relationship between the speed of the train and the level of noise emission. It also shows that at different speeds different sources of noise dominate. When stationary and at speeds below 50 km/h,

engine noise is the main source of noise from a train. If the speed is between 50 and 300 km/h rolling noise becomes the most important noise source, while at a speed exceeding 300 km/h, rolling noise is increasingly dominated by aerodynamic noise.



**Figure 7.2:** The level of different types of rail noise for different train speeds. Source: de Jong (2001).

Train speed usually lies between 50 and 300 km/h, which indicates that rolling noise (and to a lesser degree engine noise) causes the most noise disturbance. Reduction measures for noise emissions should therefore mainly focus on keeping rail infrastructure and wheels smooth and flat by frequent filing and replacement of the current block brakes by more wheel-friendly brakes. Rolling noise can also be reduced by the construction of on-vehicle noise screening systems.

Measures to reduce the curving noise include the construction of wider curves, the use of guidable wheels and better lubrication of specific parts of the wheel set. Noise generated by diesel engines can be reduced by using adequate exhaust conduit silencing and a proper positioning and embedding of the engine. Braking noise can be primarily reduced through an appropriate choice of material.

## 7.4.2 Noise Dispersion

Theoretically, the computation of the dispersion of sound from an emission point is straightforward. The noise immission level, measured as the sonic pressure, for any given point-location is a logarithmic function of the distance from and the noise level at the noise source. As the distance from the source is doubled, the decline in the noise level is about 6



dB(A)<sup>56</sup> for a point source and about 3 dB(A) for a line source (de Jong, 2001). For a railroad track with relatively little traffic, the decline in the noise level lies between that of a point source and a line source.

However, this simple relationship between noise emission and immission is affected by factors including natural landscape, morphological characteristics, meteorological conditions, noise barriers and interference of other noise sources. To reduce the immission value for a given level of noise emission, artificial noise barriers can be constructed. Noise barriers are very effective due to the fact that rolling noise is generated at a very low surface level.

## 7.5 Disturbance From Rail Noise

Although in most developed countries the population disturbed by railway traffic noise is considerably smaller than that disturbed by road traffic or aviation, it is an increasing problem. Table 7.2 shows, for various noise sources, the proportion of the Dutch population that suffers from noise nuisance for the period 1990-2002. It is interesting to see that, for rail traffic, the proportion suffering from noise has increased during this period while for the other noise sources, including road and air traffic, it has decreased.

**Table 7.2:** The percentage of the Dutch population that suffers from noise disturbance from road traffic, rail traffic and aviation between 1990 and 2002

Noise source	1990	1995	1998	1999	2000	2001	2002
Road traffic	34	30	28	28	29	30	32
Rail traffic	5	5	5	6	7	6	7
Air traffic	25	21	19	18	19	18	19

Source: CBS (2003)

Table 7.3 shows the percentage of the population in the Netherlands that disturbed by noise from various transport modes and for various noise levels. The table indicates that 6.1 per cent of the Dutch population suffers from rail noise; for noise from road traffic and aviation, this percentage is much higher. However, an interesting result in the table is that the *relative* importance of railway traffic versus that of road and aviation transport increases as the noise level increases. This means that at higher dB(A) levels, rail noise is more likely to result in noise disturbance than noise from other transport modes. An explanation for this can be found

<sup>56</sup> dB(A) or decibel is a widely used measure of the loudness of sound.

in a study by Scholten et al. (2000), who show that residential densities in the Netherlands tends to be high around railway stations and that areas close to railway tracks are intensively used for residential construction due to lack of space.

**Table 7.3:** The percentage of the Dutch population that suffers from noise disturbance from road traffic, rail traffic and aviation

Noise source	55-60 dB(A)	60-65	65-70	70-75	>75	Total
Road traffic	34.0	16.0	2.7	1.0	0.3	54.0
Rail traffic	4.5	0.9	0.3	0.2	0.13	6.1
Aviation	21.0	12.0	2.0	0.7	0.3	36.0

Source: IWW/Infras (1995)

The effect of habituation to rail noise on the degree of disturbance is investigated in a Dutch study by Dongen et al. (1982). This study compares the disturbance caused by a newly operational railroad line at two different moments; three and twenty-one months after the line became operational. We used data from this study to conduct an ordered probit analysis on the effects of the habituation to noise on the degree of noise disturbance, controlling for noise level<sup>57</sup>. The results in Table 7.4 show that as people become accustomed to rail noise exposure, the degree of disturbance decreases. Furthermore, the coefficient for noise level shows that there is a positive and significant relationship between noise level and the degree of noise disturbance.

**Table 7.4:** The effect of habituation to noise on the degree of disturbance, controlling for noise level.

	Coefficient	z-value
dB(A)	0.064	5.100
Habituation to noise	-0.223	-2.670

N = 266

Pseudo R-square = 0.044

Source: estimates are based on micro data in Dongen et al. (1982).

Table 7.5 shows the results of an ordered probit analysis, based on data from the same study,<sup>58</sup> on the effect of the usage of a train on the perceived disturbance. The results show

<sup>57</sup> The degree of annoyance consists of four categories: not aware of the noise, not disturbed, disturbed and seriously disturbed. Noise level is a continuous variable, measured in dB(A). Habituation is measured by means of a dummy which has value 0 for observations shortly after the opening of the line and value 1½ years later.

<sup>58</sup> The degree of annoyance and the noise level are measured in the same way as in the previously mentioned probit model. The usage dummy has value 1 if a person uses the railroad line for transportation purposes and value 0 if he or she does not use the railroad line.

that the group of people that uses the railroad track are less disturbed than the non-user group. Also here, we controlled for the noise level. The coefficient of the noise level is again positive.

**Table 7.5:** The effect of train use on the degree of rail noise disturbance, controlling for noise level.

	Coefficient	z-value
dB(A)	0.044	2.890
Usage dummy	-0.556	-3.170
N = 266		
Pseudo R-square = 0.043		

Source: estimates are based on micro data in Dongen et al. (1982).

Further research (see Peeters et al., 1982) shows that noise from rail traffic causes less general, non-specific disturbance, but is more disturbing than noise from road traffic when listening to television or radio, during conversations and when performing tasks that demand concentration. The most disturbing elements of rail noise are passing freight trains, work on the line, and traffic signals. Further results show that the orientation of the house with respect to the railroad track (parallel or perpendicular) and the layout of the residence are important determinants of the experienced level of disturbance. The quality of the front-side insulation has no demonstrable influence. Among people who are lightly exposed to rail noise, certain non-auditive disturbances such as perceptions of danger or unsafety, pollution, obstruction, and disturbance of the television signal are more important than noise. Individual differences in the experience of rail traffic noise are large. The relationship between the immission noise level and the degree to which this leads to disturbance is affected by the following factors: the attitude towards the railway as an environmental element, the view on the railroad track from the living room, the sensitivity to noise, the disturbance from other noise sources, and the satisfaction with the quality of the house.

## 7.6 The Economic Valuation of Rail Noise

The fact that railway traffic causes less noise than road traffic and aviation is also reflected in the costs of noise disturbance. Table 7.6, based on data from Infrac/IWW (2004), shows the annual costs of noise per transport mode for a set of 17 European countries in 2000. The total costs of noise sum up to 0.51 per cent of the total GDP in these 17 countries. The share of the

costs from rail noise is 4.7 per cent of the total noise costs. This share varies between 0.5 per cent (Norway) to 17.5 per cent (Switzerland).

**Table 7.6:** The total costs of road, rail and aviation noise for a set of 17 European countries in 2000 (x 1 million euro)

	Road	Rail (passengers)	Rail (freight)	Aviation	Total costs	Total share of GDP
EUR 17	40,411	1,354	782	3,098	45,645	0.51%
Share	88.5%	3.0%	1.7%	6.8%		

Source: Infras/IWW (2004).

There are different methods to value the costs of noise disturbance. In the valuation literature, a distinction is made between damage costs (direct and indirect) and prevention costs<sup>59</sup>. The goal of prevention measures is to reduce the damage costs, which increase more than proportionally as noise pollution increases. Therefore, prevention costs are more effective at higher noise pollution levels. An increase in prevention costs reduces the total amount of noise pollution, which in turn reduces the damage costs. Prevention measures are feasible as long as the marginal costs of prevention measures are lower than the marginal benefit (i.e. the marginal decrease in damage costs). Table 7.7 shows a categorization of economic costs of rail noise. In sections 7.6.1 to 7.6.3 we discuss the different cost categories.

**Table 7.7:** Categories of economic costs of rail noise.

Damage costs		Prevention costs
Direct	Indirect	
Reduction of "well-being" (partly reflected by property value decline)	Medical costs  Loss of productivity	Reduction of rail related noise emission  Reduction of vehicle related noise emission  Reduction of the immision of noise  Reduction of the annoyance about immitted noise

<sup>59</sup> In this study "prevention costs" refers to both abatement costs and avoidance costs

### 7.6.1 Indirect Damage Costs

Indirect damage costs include medical costs and costs due to productivity loss. Medical costs of noise pollution can be related to physical or psychiatric medical treatment. Examples are treatment related to hearing problems and psychiatric treatments. Exposure of school buildings, medical centers and residential areas to noise can affect the human capital stock, and indirectly the economy.

### 7.6.2 Direct costs: Property Value as a Proxy

Direct costs of noise include reductions in happiness. Although these reductions are hard to evaluate directly and individually in monetary terms, they do affect economic behavior. Therefore, the economic costs of reduction in happiness can be estimated indirectly by observing economic behavior.

In the literature, the hedonic pricing technique is often used for such estimations. The basic premise of the hedonic pricing method is that the price of a marketed good is related to its characteristics, or the services it provides. Therefore, we can value the noise exposure of a residential property by looking at how the price people are willing to pay for it changes when the level of noise immission changes. Many conditioning factors such as the amount of rail traffic per hour, the distance between the residential unit and the railroad track, wind direction and noise barriers affect the immission level and hence are reflected by the property price. The results of hedonic price studies are often summarized by a Noise Depreciation Sensitivity Index (NDSI). For example, an NDSI of 0.4 per cent at a threshold value of 55 dB(A) means that the percentual depreciation of property value can be expressed in terms of noise immission as:  $[\text{immission value} - 55 \text{ dB(A)}] \times 0.4 \text{ percent}$ . A drawback of this method is that residential units do not only differ in terms of noise immission, but in many other respects, for which it is not always easy to account. Even if residential units are identical, noise immission values are often correlated with factors such as the distance to public transport possibilities, numbers of cars in the neighborhood, and so on.

The method of hedonic pricing based on NDSI values has frequently been used in the context of airport noise evaluation (see for example Schipper, 2001) and road transport noise evaluation (see for example Levinson and Gillen, 1997), but in the context of rail noise it has not yet been used. Cost-benefit analyses of measures of rail noise prevention are sometimes

based on NDSI values from hedonic pricing studies on other noise sources, mainly road transport and aviation. Those NDSI values typically vary between 0.2 per cent and 1.3 per cent (see also Schipper, 1999) depending on the source. In some studies on aviation noise, values of 3.5 per cent are even mentioned.

Hedonic price studies are not always based on an NDSI method to identify the relationship between noise level and property value. The relationship can also be identified indirectly by observing the effect of railroad proximity on property value. The result can then analogously be summarized as a Proximity Depreciation Sensitivity Index (PDSI). The idea is that an increase in the distance between the residential unit and the railroad, results in a decrease in the noise level and hence in the property value. The drawbacks of NDSI studies also apply to PDSI studies. An additional disadvantage of the PDSI is that it does not take into account travel intensity or actual immission levels. The results between NDSI and PDSI studies can differ due to the fact that the relationship between the distance from the railroad and the noise level is non-linear and disturbed by numerous conditioning factors.

**Table 7.8:** The relationship between residential property value and railway proximity

<b>Data set</b>	<b>Price elasticity</b>	<b>t-value</b>	<b>Number of observations</b>
All data	0.059	2.87	2152
Distance less than 200 m	0.040	0.93	623
Distance less than 100 m	0.102	2.09	305

Source: Strand and Vagnes (2001)

Strand and Vagnes (2001) use a log-linear multiple regression function to estimate a PDSI value based on selling prices, controlling for factors such as the size and age of the residential unit. As Table 7.8 shows, they find a positive relationship between the distance to a railroad track and the price of a residential building. The price elasticity estimate (PDSI) is 0.059. Furthermore, Table 7.8 shows that for distances below 100 meters, the elasticity is much larger than for larger distances.

A related, less frequently used method to value noise immission is the contingent valuation method. Contingent valuation is based on stated preference (or willingness to pay) rather than on revealed preference (actual behavior). The advantage of this method is that it can be applied to situations without free price formation. In addition, the contingent valuation method may identify higher values that are probably closer to the consumer surplus loss, which is not revealed by the hedonic price method (see Feitelson, 1995). A disadvantage of

the contingent valuation method is that the results may be biased because it measures intentions and not actual behavior. Table 7.9, from a German study by Weinberger et al. (1991) shows the monthly willingness to pay for noise reductions for different levels of actual noise exposure. As expected, the willingness to pay is higher for larger noise reductions. The pattern illustrated in Table 7.9 is consistent with a downward sloping demand curve for silence.

**Table 7.9:** The monthly willingness to pay for different levels of noise reduction (in euros)

Actual noise level (during daytime)	60-65 dB(A)	65-75 dB(A)
Willingness to pay for 'no noise'	21.0	24.5
Willingness to pay for 'little noise'	9.2	21.0

Source: UBA (1991)

This study shows another disadvantage of using contingent valuation method. The use of questionnaires requires a categorization of noise into ranges of dB(A). This involves some arbitrary choices made by the analyst<sup>60</sup>. Furthermore, the categorization into ranges leads to a loss of information.

### 7.6.3 Prevention Costs

In Table 7.7, we distinguish between three types of prevention measures: reduction of noise emission, reduction of noise immission and reduction of noise disturbance. Prevention costs are costs related to for example the construction of noise barriers, vehicle noise control, renovation and building relocation. From an economical viewpoint, preventive measures should be carried out up to the point where the marginal costs of prevention become higher than the marginal damage costs. In practice, political interests may interfere with economic principles. For example, prevention costs, necessary to comply with noise emission standards, may exceed the damage costs that would occur in the absence of the emission standards, which is suboptimal from an economical point of view.

Economic valuation of noise disturbance implies that the negative effects of noise are expressed in monetary terms. Often, noise disturbance are valued indirectly, based on prevention costs instead of on damage costs. A drawback of this method is that cost calculation heavily depends on the noise limit values instituted by the government. The data in Table 7.10, from a German study by Weinberger et al. (1991) clearly show this. A lower,

<sup>60</sup> Interviews can be complemented with audio support to present noise levels in an objective way.

stricter, limit value results in a higher number of ‘overexposed’ persons, which in turn leads to an increase in abatement costs that are needed to comply with the limit values.

**Table 7.10:** The costs of noise screens for various noise limit values

Limit value (day/night)	Number of persons 'overexposed'	Costs per person in euro's	Total costs of complying with the limit values (x 1 billion euro)
70/60 dB(A)	0.67	1,620	1.09
75/65 dB(A)	1.95	1,325	2.58

Source: UBA (1991)

A somewhat different approach is taken in a study by Tyssen (1982). This study first estimates the consequences of different limit values of rail noise for existing housing construction plans. The estimations were repeated on the premise that protective noise barriers would be constructed, and the costs of such barriers were estimated. The results of this study are shown in Table 7.11. In a situation where less stringent limit values apply, the number of planned residential units that require additional noise measures is lower. Hence, the total costs of the noise barriers needed to build these planned units are lower. From these results, an implicit economic valuation of noise can be derived. A noise limit of 60 dB(A) is, at 14.5 million dollar, more than three times as expensive to maintain as a limit of 70 dB(A), which only costs 4.2 million dollar. Furthermore, after decreasing the noise level to 70 dB(A) it would cost 11.3 million dollar to decrease the noise level with an additional 10 dB(A). In other words, when a noise limit value of 60 dB(A) applies, an increase in the noise level from 60 dB(A) to 70 dB(A) leads to an implicit noise prevention cost of 11.3 million dollar, or about 1.1 million per dB(A).

**Table 7.11:** The level of prevention costs for different noise limit values

Limit value	Number of residential units that require noise reduction measures	Costs of screens (x 1 million \$)	Costs per residential unit
60 dB(A) overall	9465	14.5	1532
60/65 dB(A) combination	6575	11.3	1718
65 dB(A) overall	5530	9.3	1670
65/70 dB(A) combination	2405	5.0	2089
70 dB(A) overall	1910	4.2	2200

Source: Tyssen (1982)



An interesting observation that follows from the last column in Table 7.11 is that the cost effectiveness of the construction of noise barriers is higher in situations where more stringent limit values apply. This observation is also consistent with the result in table 7.10. The reason is that the number of buildings, which are planned but cannot be constructed in the absence of noise barriers, is higher in situations with more stringent limit values.

A study by Oertli and Wassmer (1996) on the cost-effectiveness of noise barriers on a takes a somewhat different approach. They assume a fixed budget and calculate a cost-benefit index for four different scenarios, in which they look at the decrease in dB(A) on a specific railroad track and the number of people ( $N$ ) that actually benefit from the noise barriers. They calculate the cost-benefit index (CBI) as:

$$\text{CBI} = \frac{\text{yearly costs}}{N [\text{dB(A)}_{\text{old}} - \text{dB(A)}_{\text{new}}]} \quad (7.1)$$

The CBIs, which can be interpreted as the cost per dB(A) reduction per person, for the four different scenarios range from 17 to 142 dollars. Like the results in Table 7.11, these indices can be interpreted as economic valuations of noise prevention.

In this section, a number of studies are discussed to illustrate the different methods used in the literature to evaluate the costs of rail noise. Most of the literature on noise evaluation focuses either on prevention costs or on damage costs. The studies that focus on prevention costs (usually government research) typically report costs for various limit values, or individual costs and noise reduction for a variety of measures (see for example KPMG, 2000) without paying attention to the benefits by valuating the noise reduction. Even studies that do compare prevention costs and damage cost reduction in the form of a cost-benefit analysis usually value the noise reduction with an NDSI value found in other research on noise valuation, usually non-rail based (see for example Nijland, 2001). As such, theoretical insights (e.g., marginal cost- and benefit behavior) have not yet been properly applied to empirical research and project evaluation within the field of rail noise.

## 7.7 Conclusion

The economic valuation of road transport and aviation noise pollution has been studied extensively. Rail transport remains rather underexposed in this respect. This is primarily due

to the fact that in terms of total costs, noise pollution of rail transport is of less importance. Among a set of 17 European countries, the share of rail noise costs in total noise costs ranges from 0.5 to 17.5 percent, with an average share of 4.7 percent. In this study, a conceptual model is presented that describes the causal relationships between railway system characteristics, noise emission, noise immission, noise disturbance, and economic costs of noise. We identify several factors that influence these relationships. The relationship between noise emission and immission is disturbed by various conditioning factors that include weather conditions, natural and artificial barriers and the distance between the railroad track and the immission point. According to expectations, the rail noise immission level is positively related to the degree of noise disturbance. Furthermore, the degree of habituation to rail noise is negatively related to the degree of disturbance. Whether or not people make use of a specific railway track also affects their noise experience. Rail users exhibit a lower degree of disturbance than non-users.

An important aspect of economic valuation of noise is the interaction between prevention costs and direct damage costs of noise pollution. Noise prevention policy can be aimed at several components in the conceptual model on rail noise (e.g., emission and immission reduction). The inclusion of government policy as a component in the model establishes a feedback loop between the economic costs of noise and the intermediate components in the model.

Government policy with respect rail noise is often based on cost-benefit studies that analyze the trade-off mechanisms between direct costs and prevention costs. Cost-benefit studies on rail noise policy generally use NDSI (noise depreciation sensitivity index) values from hedonic pricing studies on noise valuation of road transport and aviation transport as input values. The implicit assumption of transferability of such index values is not completely accurate, though. Noise is a complex multi-faceted phenomenon. The social and economic consequences of noise pollution do not just depend on the noise *level* (which is hard enough to measure accurately itself), but also on noise characteristics such as the type of noise, frequency, temporal distribution and subjective characteristics including attitude, habituation, activity pattern. These factors complicate the easy transfer of NDSI values between cost-benefit studies on different transport modes. Even in the case of studies on the same mode, such value transfer should be undertaken with caution. We found only one study that estimates an NDSI based on railroad data. This study (Strand and Vagnes, 2001) used proximity to a railroad instead of noise level as the independent variable. In this study, a price

elasticity of proximity with value 0.06 is found. We also found some studies that investigate the prevention costs associated with different limit values. In both of these studies, the level of total costs is higher for lower limit values. However, the cost per person or per residential unit is lower for lower limit values.

A statistical comparative analysis on the economic valuation of rail noise is difficult to carry out, due to the fact that the number of studies we find on this subject is limited and the methods used for economic valuation in the underlying studies are heterogeneous.<sup>61</sup> A more extensive and homogeneous set of case-studies is required for successful application of statistics oriented meta-analytical methods in order to uncover useful information from the existing literature on the economic evaluation of noise pollution from rail transport. In addition, more comparative cross-section research is necessary, for instance in Europe. Such research would be particularly interesting if the same line and same train goes through cities in different countries. Thus it can be questioned whether the same train on the same rail track causes an equal level of disturbance in different countries. Furthermore, one can test whether in such a set-up the same level of disturbance leads to the same level of damage value.

---

<sup>61</sup> An overview of studies on rail noise is given in Appendix 7.

## Appendix 7 : Overview of primary studies on the economic valuation of rail noise

Study	Year	Effect size type	Effect size estimations	Evaluation method	Prevention measure	Location	Within-study variation
Ellwanger	1987	Total costs of noise in Germany per 1000 passenger kilometers	\$ 0.82	HP	NA	Germany	Passenger transport versus freight transport
Infras/IWW	1995	Costs of noise annoyance per annoyed person for a given decibel level per year	\$ 56 (55-60 dB) \$ 224 (60-65 dB) \$ 560 (65-70 dB) \$ 1118 (70-75 dB) \$ 2114 (>75dB)	HP	NA	Sweden	Noise level ranges
van Kempen	2001	Total property value depreciation due to noise in the Netherlands for a given decibel level range	\$ 652.1 mln (56-60dB) \$ 781.5 mln (61-65 dB) \$ 535.28 mln (66-70 dB) \$ 289.5 mln (71-75 dB) \$ 133.1 mln (76-80 dB) \$ 40.5 mln (>80 dB)	HP	NA	The Netherlands	Noise level ranges
		Total property value depreciation due to noise in the Netherlands for a given decibel level range	\$ 637.5 mln (56-60dB) \$ 764.1 mln (61-65 dB) \$ 535.28 mln (66-70 dB) \$ 291.7 mln (71-75 dB) \$ 130.3 mln (76-80 dB) \$ 39.2 mln (>80 dB)	CV	NA	The Netherlands	Noise level ranges
KPMG-BEA	1998	Costs per decibel reduction	Various noise source related cost drivers	AC	Various noise source related measures	The Netherlands	Prevention measures and train type
Oertli and Wassmer	1995	Costs per decibel reduction per inhabitant per year	\$ 17-142	AC	Noise barriers	Switzerland	Different scenario's
Strand and Vagnes	2001	Property value elasticity of distance	0.102 - 0.059	HP	NA	Norway	Proximity circles
Tyssen	1982	Direct costs per residential unit to comply with a given limit value	\$ 1532 (60dB) \$ 1718 (60/65 dB) \$ 1670 (65 dB) \$ 2089 (65/70 dB) \$ 2200 (70 dB)	AC	Noise barriers	The Netherlands	Different limit values of noise levels
Weinberger et al.	1991	Costs per person per year to comply with a given limit value	\$ 98.0 (75/65 dB) \$ 80.2 (70/60 dB)	AC	Noise Barriers	Germany	Different limit values of noise levels
		Willingness to pay per person per month for a given noise reduction	\$ 10.83 (60-65 dB to 'little noise') \$ 24.67 (60-65 dB to 'no noise') \$ 24.67 (65-75 to 'little noise') \$ 28.88 (65-75 to 'no noise')	CV	NA	Germany	Actual noise level and 'proposed' level



## **PART IV**

## **SYNTHESIS**



# Chapter 8

## Conclusions and Discussion

Vast improvements in living standards and transportation systems in the twentieth century have instigated a spectacular growth in the demand for motorized transportation worldwide. The availability of efficient and effective transportation networks is a crucial precondition for economic development and an asset for local, regional and international accessibility; historical evidence indicates a strong correlation between economic growth and growth in transport. However, the growth in transport has also caused an increase in adverse external effects such as air pollution, noise pollution, traffic accidents and congestion. This has caused national as well as international governmental bodies to worry about the sustainability of their transport systems. Important conditions for a sustainable transport system are an accurate valuation of external effects, well-directed effective price policy and efficient public transport systems. Within the framework of this thesis, our aim is to apply existing and newly developed meta-analytical techniques in order to explain the empirical variation in (i) the price elasticity of demand for aviation transport (ii) the price elasticity of demand for gasoline demand (iii) the technical efficiency of urban public transport and (iv) the economic costs of rail noise per noise unit. In doing so, we focus on the estimation of combined mean values for these indicators and on identification of the impact of study characteristics and other conditioning factors on the value of these indicators. A common problem in meta-analysis in economics is statistical correlation between observations from the same study. In order to ensure the use of efficient estimation techniques, we perform a Monte Carlo study to compare the performance of various estimation techniques to account for such within-study dependence. In Section 8.1 we, discuss the methodological insights of the thesis. In Section 8.2, we focus on the main empirical results. Next, section 8.3 discusses some policy implications of the results. Finally, in Section 8.4, we provide some directions for further research.



## 8.1 Methodological Insights

In Chapter 2, we present the background, advantages and methodology of meta-analysis. Meta-analysis is a research method that is based on the quantitative summarization of previously documented study results on a specific effect of interest, which is referred to as the 'effect size'. Due to the largely non-experimental set-up in economic research, the number of sources of heterogeneity among studies is greater than in the experimental sciences, which leads to systematic variation in the effect sizes. Hence, the use of multivariate meta-analytical techniques, in combination with descriptive statistics, appears to be a suitable approach to investigate the transport-economic research questions presented in the introduction. We follow this approach in three of the four empirical studies in this thesis.<sup>62</sup> We use descriptive statistics to explore the variation in effect sizes and meta-regression analysis to explain the variation in effect sizes. Meta-regression analysis enables us to control for cultural, geographic, economic and institutional factors as well as study characteristics. As such, it allows us to combine data from a wide range of sources. By pooling all the study results together, we reduce the effect of random error and hence produce more reliable and precise estimates of the effect sizes of interest, which are not constrained by the characteristics of a single empirical study. In the applied meta-analytical studies within this thesis, this results in the estimation of mean values for the price elasticity of demand for aviation transport, the price elasticity of demand for gasoline and the technical efficiency of urban public transport operators. In addition, the use of meta-regression enables us to investigate the impact of conditioning factors on the value of the effect size. In the studies on the aviation demand and demand for gasoline, this provides us with many interesting insights into consumer demand behavior. In the study on public transport efficiency, this leads to interesting conclusions on the potential benefits of a deregulated environment for urban public transport. The empirical results and policy implications of these studies, which are discussed in Section 8.3 and 8.4, demonstrate that meta-analysis is a valuable research method, which enables us to obtain many insights that would not have been possible to obtain by means of primary research.

We also point to a number of potential problems in meta-analysis: publication bias, heterogeneity of effect sizes and multiple sampling. Publication bias occurs when the probability of publication of a study is affected by certain conditioning factors so that

---

<sup>62</sup> Application in the study on the valuation of rail noise suffered from two problems. First, the number of observations was too low to perform any meaningful statistical analysis. Second, the heterogeneity in effect sizes was such that it rendered the studies incomparable in a quantitative analysis.

published studies may not be representative of the population of studies. Detection methods and remedies for publication bias have been developed but the validity of these methods for the typical set-up of meta-analysis in economics is uncertain. More research is required on how publication bias might work in a set-up where the set of effect sizes is based on multiple sampling techniques and where there is systematic variation in the effect sizes before such methods should be used in meta-analysis in economics.

The large amount of heterogeneity in economic results leads to systematic variation in effect sizes, which can be accounted for by using a meta-regression model. The choice between a fixed effects regression model and a mixed effects model is commonly based on a homogeneity test on the effect size variation that remains after correcting for the systematic variation. However, in a dataset based on multiple sampling, any remaining variation in effect sizes might actually be caused by within-study dependency. Homogeneity tests such as the Q-test do not distinguish between within-study dependency and random heterogeneity between effect sizes and may therefore support the application of inefficient estimation methods that account for heterogeneity but do not take into account the dependency structure in the dataset.

While no estimation techniques have been designed to deal specifically with dependence in meta-analysis, both the field of spatial econometrics and hierarchical modeling provide statistical tests for dependence between observations, as well as estimators that account for such dependence. By means of a Monte Carlo study, described in Chapter 3, we compare the performance of four estimation methods, including OLS and a 2-level variance components estimator in the case of within-study dependence. The results show that estimation techniques that account for dependence between observations perform better than OLS. The 2-level estimator performs best in handling the dependence structure that pervades meta-analysis. Based on the results of this analysis, we advocate (i) the use of a statistical test based on the multilevel estimator in order to test for within-study dependence and, if such dependence is present, (ii) the use of the multilevel model for estimation in applied meta-analytical studies. We follow this approach ourselves in the applied meta-analytical studies in this thesis. In the case of the studies on aviation demand and technical efficiency of urban public transport, statistical testing indicates the presence of within-study dependency in the dataset. Hence, in these studies we use the multilevel model for estimation.

In the study described in Chapter 5, we develop a model based on a system of meta-regression equations (SMR) which allows for the simultaneous estimation of a set of meta-regression analyses in which the effect sizes are linearly related. This model has two important advantages on commonly applied meta-regression methods. First, it increases the

amount of information that can be used for the analysis. In the case of the study in Chapter 5, the results show that the associated increase in sample size leads to lower standard errors. Second, it enables the analyst to decompose the effect of interest into partial effects and analyze the impact of conditioning factors on the effect of interest by looking at the impact on the partial effects. If certain conditions concerning the number of observations are met, the SMR model appears to be a useful estimation technique for meta-analyses in which the effect size of interest can be decomposed into partial effects that are by themselves of interest to the analyst. It appears to be a particularly useful approach if the effect size is an elasticity.

## 8.1 Results from Empirical Analysis

In Chapter 4, we describe a meta-analytical study that investigates the empirical variation in the price elasticity of demand for aviation transport. In this study, based on a set of 204 observations, we find a mean price elasticity of  $-1.146$ , which indicates that price changes result in a more than proportionate change in demand for aviation transport. Furthermore, by means of meta-regression, this study investigates the impact of several determinant factors such as distance, fare class or study characteristics on the price elasticity of demand. The most important results are as follows. Price elasticities are lower in the long run, which indicates that longer response time enables consumers to adjust better to price changes. The coefficient of the trend variable indicates that passengers have become less price sensitive over time, which could indicate increased dependency on aviation transport due to growth of the air transport sector. Further results show that business passengers are less sensitive to price; this is a common finding in the literature, which can be explained from differences in the number of substitutes and the valuation of time and quality aspects between business and leisure passengers. The difference in price elasticity is about 0.6. Despite differences in substitution possibilities, no significant differences in price elasticities are found between geographical regions. Finally, case studies that do not correct for income find lower price sensitivity, which suggests that in new case studies, income should not be left out.

In the meta-analytical study described in Chapter 5, we aim to explain the empirical variation in the price elasticity of gasoline demand of car users. The price elasticity of gasoline can be seen as the sum of three partial elasticities, i.e. the price elasticities of gasoline efficiency, car use and car ownership. We estimate weighted mean values for the price elasticities of gasoline demand and each of the partial elasticities, using the linear

relationships between these elasticities as constraints. We find that the mean price elasticity of gasoline demand is inelastic at -0.53. Apparently, automobilists are not very price sensitive. The results on the partial elasticities indicate that the response in demand for gasoline to a change in price is mostly driven by changes in fuel efficiency (0.22) and car ownership (-0.22) and to a lesser degree by a change in car use (-0.10). Next, by means of meta-regression analysis, we aim to identify the impact of determinant factors that affect the price elasticities of gasoline demand and the three partial elasticities. We estimate a model based on a system of meta-regression equations in which we use the linear relationship between the elasticities as a constraint. The most important results are as follows. Price sensitivity with respect to gasoline demand is lower in the US, Canada and Australia than in the rest of the world, which is primarily due to lower price sensitivity with respect to car ownership. The high dependence of in those countries on automobile transport can be explained from the combination of high income, underdeveloped public transport and relatively low fuel prices. Long-run price elasticities are lower than short-run elasticities. Surprisingly, this is caused by long-run adjustments in the mileage per car. Furthermore, we find that consumers have become more price sensitive over time, which might be explained by an increase in the share of income spent on gasoline. A number of other characteristics of the underlying study that affect the price elasticity are the use of a dynamic versus a static specification and the type of data (cross-section, time series etc.) that was used.

In chapter 6, we investigate the empirical variation in the technical efficiency of urban public transport operators by means of a meta-analytical study. Based on a dataset of 93 observations we find an average technical efficiency ratio of 0.825. Subsequently we use meta-regression analysis in order to identify the impact of study characteristics and conditioning factors. The most important results are as follows. The coefficients that we find for the GDP, the government share of GDP, geographic dummy variables and the time trend variable, directly or indirectly suggest a negative relationship between state regulation and technical efficiency. Hence, these results support the theory that a deregulated environment is beneficiary for the efficiency of public transport operators. Other characteristics of the underlying study that affect the technical efficiency include the sample size, the type of data used (cross-section, time series etc.), the assumptions on which the efficiency frontier is based, the number of inputs that were used and the output measure on which the estimate was based.

In Chapter 7, we describe a research survey on the external costs of noise from rail transport. In this chapter, a conceptual model is presented in which railroad traffic density,

noise emission, noise immission and noise disturbance are causally related. Noise disturbance in its turn causes social and economic costs, such as property value depreciation. We find that the economic costs of noise pollution do not only depend on the noise level but also on the type of noise, the frequency, the temporal distribution and on many subjective characteristics including attitude, habituation and activity pattern. Policy measures, aimed at reducing social and economic costs, are directed at various stages of the conceptual model. These measures can be subdivided into regulation and prevention measures. We find that stricter threshold values lead to higher total costs, but may decrease social costs per capita. The economic feasibility of policy measures is usually analyzed by means of a cost-benefit case study. However, methods of analysis used are diverse and ad hoc. Therefore, results of different case studies are not easily compared in the context of a research synthesis. The heterogeneity between case studies has led to the application of a wide range of valuation methods. In many cases, a combination of valuation methods appears to be the most optimal approach to a realistic valuation of external costs.

### **8.3 Policy Implications**

On 12 September 2001, the Commission presented its White Paper on the future common transport policy: "European transport policy for 2010: time to decide" (EC, 2001b). The main objectives of the common transport policy were: (i) having taxation systems reflect the true costs of transport, including external costs such as environmental damage, congestion, or human accidents; (ii) making transport systems more efficient and safer; (iii) shifting the balance between modes of transport by 2010 by revitalizing railways and promoting maritime and inland waterway transport.

Pricing policy is a potentially powerful economic instrument of transport policy. Due to the entry of low cost carriers, the relative share of price charges in the passenger's total costs of flight has increased in recent years. Hence, the potential effect of price charging in the aviation sector has significantly increased. The EC advocates the eventually removal by member states of fuel tax exemptions traditionally applied to the aviation sector. Currently fuel taxes are allowed for domestic flights in Member States (EC, 2003b), but are often still impossible to impose on international flights. Furthermore, the EU emphasizes the introduction of economic and regulatory market incentives to enhance the competitive edge of operators and users which choose to use state-of-the-art technologies and environmentally

friendly operations (EC, 1999). Submission to a system of limited, tradable, pollution rights is considered as the most effective economic option to direct airline operators towards more environmentally friendly operations by the EU (EC, 2005b).

In order to be able to estimate the expected effects of price policy the government needs information on the price sensitivity of passengers. In the study on the price elasticity of aviation demand, we find a mean value of the price elasticity of -1.15. This indicates that, on average, aviation passengers are reasonably sensitive to price charges; an increase in price would lead to a more than proportionate decrease in demand. The results of the meta-regression analysis indicate that the price sensitivity is much lower for business class passengers than for economy class passengers. This may give the airlines the opportunity to pass on part of the extra costs (resulting from a price-based policy instrument) to the passengers, by charging the business passengers more than proportionally. If this is not acknowledged by the authority, the (environmental) policy may be less successful. However, in general the results suggest that price charging is an effective instrument to direct airlines towards more environment-friendly policies; additionally, the revenues from the charges can be used to finance charging systems which focus on price charges in combination with the subsidization of environment-friendly aviation technologies. Current EU policy on fuel taxes for domestic flights and proposals on a further decrease of tax exemptions for aviation fuel and the promotion of environment-friendly technologies and operations could support the development and application of such charging systems.

A wide range of fiscal instruments is applied to the use of road transport in the European Union, including registration tax, road and insurance tax, fuel taxes and infrastructure charges. Discussion of economic instruments focuses on methods of improving the tax system to better align usage costs with external effects rather than determining the optimal tax. Recent political activities provide the impression that European economies will face in the long run a change in the structure of transport cost towards variabilization strategies, which shift taxation away from fixed fees such as vehicle licensing to usage fees such as fuel taxes and road usage pricing (Kariske, 2003; English et al., 2000). Recent European fuel pricing policy mainly focuses on harmonization of member states tax rates and on the promotion of the use of bio-fuels or other renewable fuels for transport (EC, 2003a,b); minimum rates of taxation are set for *inter alia* motor fuel although member states are allowed to apply total or partial exemptions or reductions in the level of taxation to, *inter alia* biofuels. Efforts to promote fuel efficiency in the European Union are mainly represented by regulations on pollution emissions and consumer information on fuel economy. In the US,

mandatory corporate fuel efficiency (CAFE) standards have been in effect since 1975; car manufacturers who fail to meet the standards are confronted with financial penalties.<sup>63</sup> There has been considerable debate about the influence of the standards. Proponents see them as a proven<sup>64</sup> way to decrease the demand for gasoline while opponents argue that they are a costly and cumbersome way to reduce gasoline consumption and that fuel taxes are more effective. The results of the meta-analysis discussed in Chapter 5 indicate that, with a mean price elasticity of -0.53, the demand for gasoline is not very price sensitive. This lends some support to the view that policy based on fuel efficiency standards such as CAFE might be a more effective way to reduce gasoline demand. Further results show that the demand response to a change in the fuel price is mainly caused by responses in fuel efficiency and car ownership, which indicates that a price charge would have a larger impact on the level of energy efficiency and the number of cars owned than on the mileage of cars. The relatively low price sensitivity of car use, compared to that of car ownership, casts some doubts on the expected benefits of a shift from vehicle-based taxes to mileage-based taxes and points out that the effects on car ownership should not be ignored in the discussion on variabilization. The low price sensitivity in general indicates that fuel taxes alone do not seem to be very effective in reducing the external costs of road transport and that additional tax instruments are necessary. Taxes on purchasing prices, registration fees and taxes on fuels are in force in most member states but, according to analysts, the current structure and level of these charges are a disincentive to energy efficiency. This suggests that the EU should focus on the possibilities of actively promoting smart charging systems that stimulate the use of energy efficient technology through fuel efficiency tax-cuts and subsidization. In this context, one could think of innovative tax arrangements such as the "feebate" concept, based on taxes for vehicles with high fuel consumption along with rebates for vehicles with low fuel consumption, such as introduced in certain countries.

For several decades in Western Europe, regulation of local public transport effectively ruled out competition. In the last 10 years, the economic conditions of the public transport sector in the EU have changed a great deal. Most member states have now introduced an element of competition in their legislation or administrative practices, relating to at least part of their public transport market. In nearly all of these cases 'controlled competition' is used - based on the regular renewal of exclusive rights, rather than on free access to the market (Hidson and Müller). Currently, as implemented by Regulations No 1107/70 and No 1191/69,

---

<sup>63</sup> As of 1990, the CAFE standard for passenger car fuel economy is 11.7 kilometer per liter.

<sup>64</sup> See for example Goldberg (1998)

only in air or maritime transport European Community legislation requires competitive tendering for public service contracts. In inland transport, it is restricted to imposing strict methods of calculation for the amount of public service compensation. Recently, the European Union designed a proposal (EC, 2005a) based on the following principles: (i) contractual arrangements between the competent authority and the operator(s) responsible for providing services where financial compensation or exclusive rights are granted; (ii) need for a periodic review of contract terms; (iii) competitive tendering for public service contracts, except in certain cases where direct award is possible. The results of the meta-analysis discussed in Chapter 6 suggest that a deregulated environment is beneficiary for the efficiency of public transport operators. The proposal of the EU, currently under consultation, is a step towards a further reduction of the role of the state. While the benefits of the principles on which the proposal is based are supported by the results of the meta-analysis, the consultation should be supported by dedicated research that also pays attention to the (negative) effects of deregulation on other aspects of public transport such as the service frequency, coverage, ticket price and quality aspects.

European policy on noise from rail transport is reflected by legislation on the interoperability of the trans-European rail system, in which requirements are set for the technical specifications of rolling stock and, in the case of the high-speed rail system, railway infrastructure (EC, 1996, 2001a, 2002a, 2002b, 2004, 2006). On a state level, government policy is often based on cost-benefit studies that analyze the trade-off between damage costs and prevention costs. These studies commonly use noise depreciation sensitivity indices (NDSI) values from hedonic pricing studies on noise from road transport and aviation. In the review study on rail noise, we find that the economic costs of noise pollution do not depend only on the noise level but also on many other factors, which complicates the transfer of NDSI values to cost-benefit studies on different transport modes and may lead to under- or overestimation of the true costs and thus to suboptimal policy decisions. Ideally, policy should be based on NDSI values from rail noise studies, which account for case-specific characteristics and conditioning factors. However, in order to obtain such NDSI-values, more primary research is needed.



## 8.4 Directions for Further Research

Both the applied and the methodological research in this thesis raise new research questions. First, the application of research synthesis, although a very useful technique, does not take away the need for primary research. Meta-analysis can only be carried out on the basis of a sufficient number of similar case studies. This is notable in the study on rail noise, in which we find that the monetary valuation of rail noise has not received much attention in the research literature. The fact that cost-benefit analysis on rail noise prevention measures is often based on NDSI values from hedonic pricing studies on road transport and aviation, serves as a clear indication that more primary research is needed on the estimation of NDSI values of rail noise. While the literature on price elasticities of gasoline demand is quite large, the number of reported observations on gasoline price elasticities with respect to car use, car ownership and fuel efficiency is rather small. More research on the impact on gasoline prices on each of these variables is required for a realistic evaluation of the true effects of pricing policy. The aviation market continues to evolve rapidly. News about bankruptcies, alliances and entries of new low-cost carriers still reaches us on a regular basis. Hence, in the near future, it might be a good idea to re-estimate the meta-analysis discussed in Chapter 4, with new observations added, and to compare the results. Due to the increasing importance of low-cost carriers, it might be useful to include the absolute price level as an explanatory variable. The results of the study on public transport efficiency suggest a negative relationship between state regulation and technical efficiency. This conclusion requires confirmation and further investigation by a state-of-the-art study in which a large set of technical efficiency observations from as many countries as possible is regressed on government expenditure share or, if available, data on public transport subsidies. The advantage of such a large dataset is that all observations are compared to the same efficiency frontier. Furthermore, all sources of heterogeneity due to study characteristics are removed. In general, as follow-ups on the applied meta-regression analyses in this thesis, state-of-the art primary analyses could be conducted, based on the insights offered by the results of the meta-analyses. The results of these analyses could be compared with the conditional results predicted by the meta-analyses. Similar results would indicate that the meta-models are useful instruments for the purpose of benefit transfer and/or conditional forecasting.

The design of the simulation study described in Chapter 3 is based on a number of assumptions about homoskedasticity and the structure and degree of dependence between

observations. Due to differences in sample sizes in applied meta-analytical datasets, the assumption of homoskedasticity might not conform to reality. Hence, it would be interesting to see how the results are effected if we allow for heteroskedasticity in the simulation. Likewise, the assumption of equal degree of within-study correlation among studies may not hold in applied meta-analysis; one might expect a lower degree of within-study dependence for heterogeneous studies than for homogeneous studies. To account for this, one could allow for different levels of within-study correlation, which might for example be based on some indicator of the level of heterogeneity within a study. The simulation design is furthermore based on the notion that multiple sampling causes statistical correlation between observations from the same study, as such observations are generated under similar circumstances. However, there may also be statistical correlation between observations that come from different studies with very similar characteristics. This could be accounted for by generating a dependence structure in which the correlation between any two observations is based on, for example, the amount of study characteristics that they have in common. Finally, we assume that the within-study dependence pervades both the observed and the unobserved part of the variation to the same degree. While this assumption is based on realistic expectations about the pattern of dependence, it would be interesting to analyze the case with only dependence in the unobserved part.



# References

- Abdel-Khalek, G. (1988), Income and price elasticities of energy consumption in Egypt: a time-series analysis, *Energy Economics*, vol. 10, pp. 47-58.
- Abrahams, A. (1983), A service model of air travel demand: an empirical study, *Transportation Research A*, vol. 17, pp. 385-93.
- Agarwal, V. and W.K. Talley, (1985), The demand for international air passenger service provided by U.S. air carriers, *International Journal of Transport Economics*, vol.12, pp. 63-70.
- Alperovich, G. and Y. Machnes (1994), The role of wealth in the demand for international air travel, *Journal of Transport Economics and Policy*, vol. 28, pp. 163-173.
- Andrikopoulos, A.A. and T. Teriovitis (1983), An abstract mode model: a cross-section and time series investigation, *International Journal of Transport Economics*, vol. 10, pp. 547-555.
- Anselin, L. (1988), *Spatial econometrics: methods and models*, Kluwer, Dordrecht.
- Anselin, L. and S. Hudak (1992), Spatial econometrics in practice: a review of software options, *Regional Science and Urban Economics*, vol. 22, pp. 509-536.
- Anselin, L. and W.K. Tam Cho (2002), Spatial Effects and Ecological Inference, *Political Analysis*, vol. 10, pp. 276-297.
- Archibald, R. and R. Gillingham (1980), An analysis of the short-run consumer demand for gasoline using household survey data, *Review of Economics and Statistics*, vol. 62, pp. 622-628.
- Archibald R and R. Gillingham (1981a), A decomposition of the price and income elasticities of the consumer demand for gasoline, *Southern Economic Journal*, vol. 47, pp. 1021-1031.
- Archibald, R. and R. Gillingham (1981b), The distributional impact of alternative gasoline conservation policies, *The Bell Journal of Economics*, vol. 12, pp. 426-444.
- Aschenfelter, O., C. Harmon and H. Oosterbeek (1999), A review of estimates of the schooling/earnings relationship, with tests for publication bias, *Labour Economics*, vol. 6, pp. 453-470.
- Aubree, D. (1975), *La Gêne due a Bruit des Trains*, Centre Scientifique et Technique du Bâtiment (C.S.T.B.), Nantes.
- Bal, F., K.J. Button and P. Nijkamp (2002), Ceteris paribus, meta-analysis and value transfer, *Socio-Economic Planning Sciences*, vol. 36, pp. 127-138.

- Baltagi, B. and J. Griffin (1983), Gasoline demand in the OECD: an application of pooling and testing procedures, *European Economic Review*, vol. 22, pp. 117-137.
- Baltagi, B. and J. Griffin (1997), Pooled estimators vs their heterogeneous counterparts in the context of dynamic demand for gasoline, *Journal of Econometrics*, vol. 77, pp. 303-327.
- Banaszak S, U. Chakravorty and P.S. Leung (1999), Demand for ground transportation fuel and pricing policy in Asian tigers: a comparative study of Korea and Taiwan, *The Energy Journal*, vol. 20, pp. 145-165.
- Banister, D. (ed.) (1999), *Transport policy and the environment*, KY: E & F N Spon, Florence.
- Banister, D. (2000), Sustainable urban development and transport – A Eurovision for 2020, *Transport Reviews*, vol. 20, pp. 113-130.
- Begg, C.B. and J.A. Berlin (1989), Publication bias and dissemination of clinical research, *Journal of National Cancer Institute*, vol. 81, pp. 867-872.
- Bentzen, J. (1994), An empirical analysis of gasoline demand in Denmark using cointegration techniques, *Energy Economics*, vol. 16, pp. 139-143.
- Berndt, E. and G. Botero (1985), Energy demand in the transportation sector of Mexico, *Journal of Development Economics*, vol. 17, pp. 219-238.
- Bhattacharyya, A., S. Kumbhakar and A. Bhattacharyya (1995), Ownership structure and cost efficiency: a study of publicly owned passenger-bus transportation companies in India, *Journal of Productivity Analysis*, vol. 6, pp. 47-61.
- Bijmolt, T.H.A. and R.G.M. Pieters (2001), Meta-analysis in marketing when studies contain multiple measurements, *Marketing Letters*, vol. 12, p157-169.
- Blaeijs, A. de, R.J.G.M. Florax, P. Rietveld and E. Verhoef (2003), The value of statistical life in road safety: a meta-analysis, *Accident Analysis and Prevention*, vol. 35, pp. 973-986.
- Blair R, D. Kaserman and R. Tepel (1984), The impact of improved mileage on gasoline consumption, *Economic Inquiry*, vol. 22, pp. 209-217.
- Bordley, R. (1991), An overlapping choice set model of automotive price elasticities. *Transportation Research B*, vol. 28, pp. 401-407.
- Brander, L.M., R.J.G.M. Florax and J.E. Vermaat (2003), *The empirics of wetland valuation: a comprehensive summary and meta-analysis of the literature*. IVM working paper W03-30, Vrije Universiteit Amsterdam.
- Brons, M.R.E, P. Nijkamp, A.J.H. Pels and P. Rietveld (2002), Price Elasticities of Demand for Passenger Air Travel, A Meta-Analysis, *Journal of Air Transport Management*, Vol. 8, no.34, pp. 165-175.
- Brons, M.R.E, P. Nijkamp, A.J.H. Pels and P. Rietveld (2003), Rail noise: Economic Valuation and Policy. *Transportation Research D: Transport and Environment*, Vol. 8, no. 3, pp. 169-184.

- Brons, M.R.E., P. Nijkamp, A.J.H. Pels and P. Rietveld (2005), Efficiency of Urban Public Transit: a Meta-analysis, *Transportation*, Vol. 32, no. 1, pp. 1-21.
- Brouwer, R., I. Langford, I. Bateman and R.K. Turner (1999), A meta-analysis of wetland contingent valuation studies, *Regional Environmental Change*, vol. 1, pp. 47-57.
- Brown, S.L. and W.S. Watkins (1974), *The demand for air travel: a regression study of time series and cross-sectional data in the U.S. domestic market*, paper presented at the 47th annual meeting of the Civil Aeronautics Board, 1974.
- Bureau of Transport and Communications Economics (1986), *Demand for Australian domestic aviation services: forecasts by market segment*, Occasional paper 79, AGOS, Canberra.
- Bureau of Transport and Communications Economics (1988), *Trends and prospects for Australian international air transport*, Occasional paper 88, AGOS, Canberra.
- Bureau of Transport and Communications Economics (1995), *Demand elasticities for air travel to and from Australia*, Working Paper 20, BTCE, Canberra.
- Button, K (1995), *Meta-analysis and the implications of transport*, Discussion Paper TI 2005-145, Tinbergen Institute, Amsterdam.
- Button, K. and A. Costa (1999), Economic efficiency gains from urban public transport regulatory reform: two case studies of changes in Europe, *Annals of Regional Science*, vol. 33, pp. 425-438.
- Button, K. and J. Kerr (1996), Effectiveness of traffic restraint policies: a simple meta-regression analysis, *International Journal of Transport Economics*, vol. 23, pp. 213-225.
- Card, D. and A.B. Krueger (1995), Time-series minimum-wage studies: a meta-analysis, *American Economic Review*, vol. 85, pp. 238-243.
- CEI/UNEP/OECD/Austrian Federal Ministry for Environment, Youth and Family (1999), *Towards sustainable transport in the CEI countries*.
- Centre for Sustainable Transportation (2002), *Definition and Vision of Sustainable Transportation*, [www.cstctd.org](http://www.cstctd.org).
- Chang, K.-P. and P.-H. Kao (1992), The relative efficiency of public versus private municipal bus firms: an application of data envelopment analysis, *Journal of Productivity Analysis*, vol. 3, pp. 67-84.
- Charnes, A., W.W. Cooper and E. Rhodes (1978), Measuring the efficiency of decision making units, *European Journal of Operations Research*, vol. 2, pp. 429-444.
- Chu, X., G. Fielding and B. Lamar (1992), Measuring transit performance using data envelopment analysis, *Transportation Research A*, vol. 26, pp. 223-230.
- Coelli, T.J., D.S. Prasada Rao and G.E. Battese (1998), *An introduction to efficiency and productivity analysis*, Kluwer Academic Publishers, Boston.

- Cook, T., H. Cooper, D. Cordray, H. Hartman, L. Hedges, T. Louis and F. Mosteller (1992), *Meta-analysis for explanation: a casebook*, The Russell Sage Foundation Press, New York.
- Cooper, H. (1989), *Synthesizing Research. A guide for literature reviews*, Applied Social Research Methods Series vol. 2., Sage Publications, Thousand Oaks.
- Cooper, H. and Hedges, L.V.(Eds.)(1994), *The handbook of research synthesis*, Russell Sage Foundation, New York.
- Cooper, M.H. and A.K. (1971), *The price of air travel*, Hobart Paper 53, Institute for Economic Affairs, London.
- Costa, Á. (1998), Public transport efficiency and effectiveness: metro de Madrid, in K. Button, P. Nijkamp and H. Priemus (eds), *Transport Networks in Europe: Concepts, Analysis and Policies*, Elgar, Cheltenham, pp.248-264.
- Costa, Á. and R.N. Markellos (1997), Evaluating public transport efficiency with neural network models, *Transportation Research C*, vol. 5, pp. 301-312.
- Cowie, J. and D. Asenova (1999), Organization form, scale effects and efficiency in the British bus industry, *Transportation*, vol. 26, pp. 231-248.
- Crouch, G.I. (1991), *The effects of income and price on international tourism demand*, Management Paper 38, Monash University, Victoria.
- Dahl, C. (1978), American energy consumption extravagant or economical: a case study of gasoline demand, *Resources and Energy*, vol. 1, pp. 359-373.
- Dahl, C. (1979), Consumer adjustment to a gasoline tax, *Review of Economics and Statistics*, vol. 61, pp. 427-432.
- Dalhuisen, J.M., R.J.G.M. Florax, H.L.F. de Groot and P. Nijkamp (2003), Price and income elasticities of residential water demand: Why empirical estimates differ, *Land Economics*, vol. 79, pp. 292-308.
- Dassen, A.G.M., J.H.J. Dolmans, J. Jabben, N.A.R. Hamminga and H.A. Nijland (2000), *Geluid in de vijfde milieuverkenning*, RIVM rapport 408129009, Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven
- Davidse, R, C. van Driel and C. Goldenbeld (2004), *The effect of altered road markings on speed and lateral position: a meta-analysis*, SWOV Institute for Road Safety Research, Leidschendam
- De Borger, B., K. Kerstens, and Á. Costa (2002), Public transit performance: what does one learn from frontier studies?, *Transport Reviews*, vol. 22, pp. 1-38.
- De Jong, G. and F. Cheung (1999), Stochastic frontier models for public transport, in H. Meersman, E. van de Voorde and W. Winkelmans (eds), *World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research*, vol. 1: *Transport Modes and Systems*, Pergamon, New York, pp. 373-386.

- De Vaney, A.S. (1974), *The measuremental cost of airport noise*, Texas A. & M. University, Kingsville.
- Delhause, B., S. Perelman and B. Thiry (1992), Substituabilité partielle des facteurs et efficacité-coût: l'exemple des transports urbain et vicinal Belges, *Economie et Prévision*, vol. 32, pp. 105-115.
- Deprins, D., L. Simar and H. Tulkens (1984), Measuring labor efficiency in post offices, in *The Performance of Public Enterprises: Concepts and Measurement*, Elsevier Science, Amsterdam, pp. 234-267.
- Desvousges, W.H., M.C. Naughton and G.R. Parsons (1992), Benefit transfer: conceptual problems in estimating water quality benefits using existing studies, *Water Resources Research*, Vol. 28, pp. 675-683.
- Dickersin, K., Y.I. Min and C.L. Meinert (1992), Factors influencing publication of research results: follow-up of applications submitted to two institutional review boards, *Journal of the American Medical Association*, vol. 263, pp. 374-378.
- Dings, J.M.W., P. Janse, B.A. Leurs and M.D. Davidson (1999), *Efficiënte prijzen voor het verkeer. raming van maatschappelijke kosten van verschillende vervoermiddelen*, Centrum voor energiebesparing en schone technologie, Delft.
- Dongen, J.E.F., R.G. de Jong and R. van den Berg (1982), *De gewinning aan het geluid van een nieuwe spoorlijn te Zoetermeer*, IMG-rapport D50, Instituut Maatschappelijke Gezondheidszorg, Rotterdam.
- Donnelly, W. (1982), The regional demand for petrol in Australia, *Economic Record*, vol. 58, pp. 317-327.
- Doucoulagos, C. (1995), Worker participation and productivity in labor-managed and participatory capitalist firms: A meta-analysis, *Industrial and Labor Relations Review*, vol. 4, pp. 58-77.
- Drollas, L. (1984), The demand for gasoline: further evidence, *Energy Economics*, vol. 6, pp. 71-82.
- Easterly, W. and S. Rebelo (1993), Fiscal policy and economic growth, *Journal of Monetary Economics*, vol. 3(32), pp. 417-458.
- EEA, European Environment Agency (2000a), *Environmental Signal 2000*.
- EEA, European Environment Agency (2000b), *Are we moving in the right direction? Indicators in transport and environment integration in the EU*. TERM 2000.
- Egmond, P. van, P. Nijkamp and G. Vindigni (2001), *A comparative analysis of the performance of urban public transport systems in Europe*, Research memorandum 2001-39, Free University of Amsterdam.
- EIA. (2004), *Long-term historical statistics all in one place*, Annual Energy Review, Report No. DOE/EIA-0384 (2003),



- Eltony, M. (1993), Transport gasoline demand in Canada, *Journal of Transport Economics and Policy*, vol. 27, pp. 193-208.
- Eltony, M. and N. Al-Mutairi (1995), Demand for gasoline in Kuwait: an empirical analysis using cointegration techniques, *Energy Economics*, vol. 17, pp. 249-253.
- English, G., C. Schwier, R. Lake, R. Hirshhorn and R. Barton (2000), *Internalizing the social costs of transportation*, Report prepared for Transport Canada. Research and Traffic group.
- Eriksen, K. S. (2000), Calculating External costs of Transportation in Norway: Principles and Results, *European Journal of Transport and Infrastructure Research*, Special Issue: Sustainable Transport, pp. 9-25.
- Espey, M. (1998), Gasoline demand revisited: an international meta-analysis of elasticities, *Energy Economics*, vol. 20, pp. 273-295.
- Espey, M. and D.D. Thilmany (2000), Farm labor demand: A meta-regression analysis of wage elasticities, *Journal of Agricultural and Resource Economics*, vol. 25, pp. 252-266.
- Espey, M. and H. Kaufman (2000), The impact of airport noise and proximity on residential property values, *Growth and Change*, vol. 31, pp. 341-352.
- Espey, M., J. Espey and W.D. Shaw (1997), Price elasticity of residential demand for water: a meta-analysis, *Water Resources Research*, vol. 33, pp. 1369-1374.
- European Commission (1996), *On the interoperability of the trans-European high-speed rail system*, Directive 1996/94/EC, Brussels.
- European Commission (1999a), *Air transport and the environment towards meeting the challenges of sustainable development*, COM (1999) 640 final, Brussels.
- European Commission (1999b), *The availability of consumer information on fuel economy and CO<sub>2</sub> emissions in respect of the marketing of new passenger cars*, Directive 1999/94/EC, Brussels
- European Commission (2001a), *On the interoperability of the trans-European conventional rail system*, Directive 2001/16/EC, Brussels.
- European Commission (2001b), *White paper. European transport policy for 2010: time to decide*, COM (2001) 370 final, Brussels.
- European Commission (2002a), *Concerning the technical specification for interoperability relating to the rolling stock subsystem of the trans-European high-speed rail system referred to in Article 6(1) of Directive 96/48/EC*, Commission Decision 2002/735/EC, Brussels.
- European Commission (2002b), *Technical specification for interoperability relating to the infrastructure subsystem*, Commission Decision 2002/732/EC, Brussels.

- European Commission (2003a), *On the promotion of the use of biofuels or other renewable fuels for transport*, Directive 2003/30/EC, Brussels
- European Commission (2003b), *Restructuring the Community framework for the taxation of energy products and electricity*, Directive 2003/96/EC, Brussels.
- European Commission (2004), *Specifying the basic parameters of the 'Noise', 'Freight Wagons' and 'Telematic applications for freight' Technical Specifications for Interoperability EC, referred to in Directive 2001/16/EC*, Commission Decision 2004/446/ Brussels.
- European Commission (2005a), *On public passenger transport services by rail and by road*, COM (2005) 319 final, Brussels.
- European Commission (2005b), *Reducing the climate change impact of aviation*, COM(2005) 459 final, Brussels.
- European Commission (2006), *Concerning the technical specification for interoperability relating to the subsystem 'rolling stock — noise' of the trans-European conventional rail system*, Commission Decision 2006/66/EC, Brussels.
- European Council (2001), *Presidency conclusions regarding sustainable development*, Gothenburg European Council (15-16 June, 2001)
- European Union (2001), *Sustainable transport in the European Union. Managing transport demand, fostering economic growth*, Background paper for the seminar organised by the Belgian presidency of the EU, 12 July 2001.
- European Union (2002), *The 6th environmental action programme (6EAP) of the European Union - legally binding commitments and obligations*, European Environmental Bureau.
- Farrell, M. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society*, vol. 120, pp. 253-281.
- Fazioli, R., M. Filippini and P. Prioni (1993), Cost-structure and efficiency of local public transport: the case of Emilia Romagna bus companies, *International Journal of Transport Economics*, vol. 20, pp. 305-324.
- Feitelson, E. (1989), A property rights approach to transportation noise: the Israeli case in perspective, *Transportation Research A*, vol. 23, pp. 349-358.
- Filippini, M., R. Maggi and P. Prioni (1992), Inefficiency in a regulated industry: the case of Swiss regional bus companies, *Annals of Public and Cooperative Economics*, vol. 63, pp. 437-455.
- Florax, R.J.G.M. (2002a), *Accounting for dependence among study results in meta-analysis: methodology and applications to the valuation and use of natural resources*, Research Memorandum 2002-5, Free University of Amsterdam.
- Florax, R.J.G.M. (2002b), Methodological pitfalls in meta-analysis: publication bias, in R.J.G.M. Florax, P.Nijkamp and K.G.Willis (eds.), *Comparative environmental economic assessment*, Edward Elgar, Cheltenham.

- Florax, R.J.G.M. (2002c), *Meta-analysis in environmental and natural resource economics: the need for strategy, tools and protocol*, Free University of Amsterdam.
- Fridstrom, L. and H. Thune-Larsen (1989), An econometric air travel demand model for the entire conventional domestic network: the case of Norway, *Transportation Research B*, vol. 23, pp. 213-23.
- Furuichi, M. and F.S. Koppelman, (1994), An analysis of air travelers' departure airport and destination choice behavior, *Transportation Research A*, vol. 28, pp. 187-195.
- Gallini, N. (1983), Demand for gasoline in Canada, *Canadian Journal of Economics*, vol. 16, pp. 299-324.
- Ganley, J. and J. Cubbin (1992), *Public sector efficiency measurement: applications of data envelopment analysis*, Elsevier, Amsterdam.
- Gately, D. (1990), The US demand for highway travel and motor fuel, *Energy Journal*, vol. 11, pp. 59-73.
- Gately, D. (1992), Imperfect price-reversibility of US gasoline demand: asymmetric responses to price increases and decreases, *Energy Journal*, vol. 13, pp. 179-207.
- Gathon, H.-J. (1989), Indicators of partial productivity and technical efficiency in the European urban transit sector, *Annals of Public and Cooperative Economics*, vol. 60, pp. 43-59.
- Glaister, S., D. Starkie and D. Thompson (1990), The assessment: economic policy for transport, *Oxford Review of Economic Policy*, vol. 6, pp. 1-21.
- Glaister, S., D. Starkie and D. Thompson (1990), The assessment: economic policy for transport, *Oxford Review of Economic Policy*, vol. 6, pp. 1-21.
- Glass, G.V. (1976), Primary, secondary and meta-analysis of research, *Educational Researcher*, vol. 5, pp. 3-8.
- Goldberg, P.K. (1998), The effects of the corporate average fuel efficiency standards in the US, *The Journal of Industrial Economics*, vol. 46, pp. 1-33.
- Goldstein, H. (1995), *Multilevel statistical models*, Edward Arnold, London.
- Görg, H. and E. Strobl (2001), Multinational companies and productivity spillovers: A meta-analysis, *Economic Journal*, vol. 111, pp. 723-740.
- Graham, D. and S. Glaister (2002), *Review of income and price elasticities of demand for road traffic*, Centre for transportation studies, Imperial College, London.
- Greene, D. (1982), State level stock system of gasoline demand, *Transportation Research Record*, vol. 801, pp. 44-51.
- Greene, D. (1990), Café or price?: an analysis of the effects of federal fuel economy regulations and gasoline price on new car MPG, 1978-89, *The Energy Journal*, vol. 11, pp. 37-57.

- Greene, W. (2003), *Econometric analysis*, Prentice Hall, Upper Saddle River.
- Griffin, J. (1979), Statistical cost analysis revisited, *The Quarterly Journal of Economics*, vol. 93, pp. 107-29.
- Groot, R.S. de, M.A. Wilson and R.M.J. Boumans (2002), A typology for the classification, description and valuation of ecosystem functions, goods and services, *Ecological Economics*, vol. 41, pp. 393-408.
- Groot, W. and H. Maassen van den Brink (2000), Overeducation in the labor market: A meta-analysis, *Economics of Education Review*, vol. 19, pp. 149-158.
- Haitovsky, Y., I. Salomon and A. Silman (1987), The economic impact of charter flights on tourism to Israel: an econometric approach, *Journal of Transport Economics and Policy*, vol. 21, pp. 111-134.
- Hanly, M., J. Dargay and P. Goodwin (2002), *Review of price elasticities in the demand for road traffic*, ESRC TSU publication 2002/13, University of London, Centre for Transport Studies.
- Hanson, C.E. (1995), Proposed new noise impact criteria for passenger rail systems in the United States, *Journal of Sound and Vibration*, vol. 193, pp. 29-34.
- Hanusch, H. and U. Cantner (1991), Produktion öffentlicher Leistungen: Effizienz und technischer Fortschritt, *Jahrbücher für Nationalökonomie und Statistik*, vol. 208, pp. 269-384.
- Hartley, P. and C. Trengrove (1994), East-west airlines' Sydney-Melbourne service: market expansion or market diversion?, *The Economic Record*, vol. 66, pp. 203-220.
- Hawkins, R. (1999), *Review of studies on external costs of noise*, DETR, London.
- Hedges, L.V. and I. Olkin (1985), *Statistical methods for meta-analysis*, Academic Press, Orlando.
- Heien, D. (1969), Income and price lags in consumer-demand analysis, *Journal of the Royal Statistical Society*, vol. 132, pp. 265-271.
- Heijungs, R. (2001), *A theory of the environment and economic systems, a unified framework for ecological economic analysis and decision-support*, Edward Elgar, Cheltenham.
- Hensher, D.A. and J.J. Louviere (1983) Identifying individual preferences for international air fares, *Journal of Transport Economics and Policy*, vol. 17, pp. 225-245.
- Hidson, M. and Müller, M. (2003), *Improving environmental and social standards in urban public transport through competitive tendering - A Good Practice Guide*, Local Governments for Sustainability (ICLEI).
- Hooper, P. (1993), *The elasticity of demand for travel: a review*, Research Report, Insitute of Transport Studies, Sydney.

- Hox, J.J. and E.D. de Leeuw (2003), *Multilevel modeling. Methodological advances, issues and applications*, Lawrence Erlbaum Associates, Mahwah.
- Huber, P. (1979), *The behavior of maximum likelihood estimates under non-standard conditions*, Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics, vol. 1, Berkeley University of California Press.
- IATA, International Air Transport Association (2000), *European Air Traffic Forecasts 1985-2015*, produced by IATA for the Air Transport Action Group
- INFRAS/IWW (2000), *External costs of transport: accident, environmental and congestion costs in Western Europe*, commissioned by UIC, Zürich/Karlsruhe/Paris.
- INFRAS/IWW (2004), *Costes externos del transporte. Estudio de actualizacion. Resumen*, Zurich/Karlsruhe.
- Ippolito, R.A. (1981), Estimating airline demand with quality of service variables, *Journal of Transport Economics and Policy*, vol.15, pp. 7-16.
- Jarrell, S.B. and T.D. Stanley (1990), A meta-analysis of the union-nonunion wage gap, *Industrial and Labor Relations Review*, vol. 44, pp. 54-67.
- Jeppesen, T., J.A. List and H. Folmer (2001), Environmental regulation and new plant location decisions: Evidence from a meta-analysis, *Journal of Regional Science*, vol. 42, pp. 19-49.
- Jong, R.G. de (ed.) (2001), *Lawaai beheersing. Handboek voor Milieubeheer*, Samsom H.D. Tjeenk Willink, Alphen a/d Rijn.
- Jud, G.D. and J. Hyman (1976), International demand for Latin American tourism, *Growth and Change*, vol. 7, pp. 25-31.
- Jung, J.M. and E.T. Fuji (1976), The price elasticity of demand for air travel, *Journal of Transport Economics and Policy*, vol. 10, pp. 257-262.
- Kasiske, F. (2003), Supply Chains Collaboration – Opportunities for Transport Optimization, in: *Bundesvereinigung Logistik e.V. (Hrsg.): 20. Deutscher Logistik-Kongress Berlin: Grenzen überwinden – Wandel gestalten, Kongressband*, Deutscher Verkehrs-Verlag, Hamburg 2003, pp. 447-452.
- Katzman, M.T. (1987), Multiattribute utility elicitation techniques and public policy: A meta-analysis of empirical applications, *Applications of Management Science*, vol. 5, pp. 237-303.
- Kayser, H. (2000), Gasoline demand and car choice: estimating gasoline demand using household information, *Energy Economics*, vol. 22, pp. 331-348.
- Keef, S.P., L.A. Roberts (2004), The Meta-analysis of partial effect sizes, *British Journal of Mathematical and Statistical Psychology*, Vol. 57, pp. 97-129.
- Kempen, E.E.M.M. van (2001), Een schatting van de baten van geluidmaatregelen, RIVM rapport 715120004, Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven.

- Kennedy, M. (1974), An economic model of the world oil market, *Bell Journal of Economics and Management Science*, vol. 5, pp. 540-577.
- Kerstens, K. (1996), Technical efficiency measurement and explanation of French urban transit companies, *Transportation Research A*, vol. 30, pp. 431-452.
- Khumbhakar, S. and A. Bhattacharyya (1996), Productivity growth in passenger-bus transportation: a heteroskedastic error component model with unbalanced panel data, *Empirical Economics*, vol. 21, pp. 557-573.
- Knight, F.H. (1924), Some fallacies in the interpretation of social cost, *Quarterly Journal of Economics*, vol. 38, pp. 592-606.
- Koetse, M.J., R.J.G.M. Florax and H.L.F. de Groot (2005), *Correcting for primary study misspecifications in meta-analysis*, Tinbergen Discussion Paper TI 2005-029/03, Tinbergen Institute, Amsterdam
- KPMG-BEA (1998), *Stillere Treinen: een Bron van Besparingen*, KPMG-BEA, Hoofddorp.
- Kremers, H., P. Rietveld and P. Nijkamp (2000), *A Meta-Analysis of Price Elasticities of Transport Demand in a General Equilibrium Framework*, Tinbergen Discussion Paper TI 2000-060/03, Tinbergen Institute, Amsterdam.
- Lambert, J., P. Champelovier and I. Vernet (1996), Disturbance from high speed train noise: a social survey, *Journal of Sound and Vibration*, vol. 193, pp. 21-28.
- Levaggi, R. (1994), Parametric and non-parametric approach to efficiency: the case of urban transport in Italy, *Studi Economici*, vol. 49, pp. 67-88.
- Liew, C.K. and C.J. Liew (1979), Airline deregulation and its impact on intercity travel demand, *Transportation Research Record*, vol. 732, pp. 34-39.
- Lijesen, M. (1998), Analyzing cost structures of public sector activities, with an application to regional public transport, in: G.B.K. de Graan and F.G. Volmer (eds), *Performance Budgeting: A Perspective on modelling and Strategic Planning in the Public Sector in Holland*, Eburon, Delft, pp.178-187.
- Lin, A.-I., E.N. Botsas and S.A. Monroe (1985), State gasoline consumption in the USA: an econometric analysis, *Energy Economics*, vol. 7, pp. 29-36.
- Lipsey, M.W. and D.B. Wilson (2001), *Practical meta-analysis*, Sage Publications, Thousand Oaks.
- List, J.A. and C.A. Gallet (2001), What experimental protocol influence disparities between actual and hypothetical stated values? Evidence from a meta-analysis, *Environmental and Resource Economics*, vol. 20, pp. 241-254.
- Litman, T. (2005), Transportation Cost and Benefit Analysis Techniques, Estimates and Implications, *Victoria Transport Policy Institute* ([www.vtppi.org](http://www.vtppi.org)).

- Loizides, I. and B. Giahalis (1995), The performance of public enterprises: a case of the Greek railway organization, *International Journal of Transport Economics*, vol. 22, pp. 283-306.
- Loomis, J.B. and D.S. White (1996), Economic benefits of rare and endangered species: Summary and meta-analysis, *Ecological Economics*, vol. 18, pp. 197-206.
- Mackie, P.J., A.S. Fowkes, M. Wardman, G. Whelan, J. Nellthorp and J. Bates (2003), *Values of travel time savings in the UK*, ITS Working Papers 561-566, Institute for Transport Studies, Leeds.
- Marin, P.L. (1995), Competition in European aviation: pricing policy and market structure, *The Journal of Industrial Economics*, vol. 43, pp. 141-159.
- Marsan, M.A. and F.P.S. Kotoris (1993), *Competition through fares and fare accesses on the air transport markets*, Discussion Paper Series 776, CEPR, London.
- Matas, A. and J.-L. Raymond (1998), Technical characteristics and efficiency of urban bus companies: the case of Spain, *Transportation*, vol. 25, pp. 243-263.
- May, A.D., S.P. Shepherd and P.M. Timms (2000), Optimal transport strategies for European cities, *Transportation*, vol. 27, pp. 285-315.
- Mayeres, I. (2003), Taxes and transport externalities, *Public Finance and Management*, vol. 3, pp. 94-116.
- McRae, Robert (1994), Gasoline demand in developing Asian countries. *Energy Journal*, vol. 15, pp. 143-156.
- Mehta, J.S., G.V.L. Narasimham and P.A.V.B. Swamy (1978), Estimation of a dynamic demand function for gasoline with different schemes of parameter estimation, *Journal of Econometrics*, vol. 7, pp. 263-269.
- Melville, J.A. (1998), An empirical model of the demand for international air travel for the Caribbean region, *International Journal of Transport Economics*, vol. 25, pp. 313-336.
- Miedema, H.M.E. (1993), *Response functions for environmental noise in residential areas*, Report 1993/1, NIPG/TNO, Leiden.
- Milloy, H.B., L. Douglas and S.M. Sullivan (1985), Market response to discount fares on selected domestic trunk routes, *Papers of the Tenth Australian Transport Research Forum*, pp. 1-18.
- Moran, P.A.P. (1948), The interpretation of statistical maps, *Journal of the Royal Statistical Society B*, vol. 10, pp. 243-51.
- Moran, P.A.P. (1950), Notes on continuous stochastic phenomena, *Biometrika*, vol. 37, pp. 17-23.
- Morrison, S. and C. Winston (1986), *The economic effects of airline deregulation*, The Brookings Institution, Washington D.C.

- Morrison, S. and C. Winston (1989), *Enhancing the performance of the deregulated air transportation system*, The Brookings Institution, Washington D.C.
- Mount, R. and H. Williams (1981) Energy conservation, motor gasoline demand and the OECD countries, *The Review of Business and Economic Research*, vol. 16, pp. 48-57.
- Mulatu, A., R.J.G.M. Florax and C.A. Withagen (2002), Environmental regulation and competitiveness: An exploratory meta-analysis, in: A. Läscher (ed.), *Empirical Modeling of the Economy and the Environment*, Springer-Verlag, Berlin.
- Mutti, J and Y. Murai (1977), Airline travel on the North Atlantic, *Journal of Transport Economics and Policy*, vol. 11, pp. 45-53.
- Nelson, J.P. (2004), Meta-analysis of airport noise and hedonic property values, *Journal of Transport Economics and Policy*, vol. 38, pp. 1-28.
- Newman, P. and J. Kenworthy (2004), Sustainability and cities. *Overcoming automobile dependence*, Island Press, Washington D.C.
- Nijkamp, P and G. Pepping (1998), Meta-analysis for explaining the variance in public transport demand elasticities in Europe, *Journal of Transportation and Statistics*, vol. 1, pp. 1-14.
- Nijkamp, P. (2005), Meta-analysis: a methodology for research synthesis, *Studia Regionalia*, Forthcoming.
- Nijkamp, P. and J. Poot (2004), Meta-analysis of the impact of fiscal policies on long-run growth, *European Journal of Political Economy*, Vol. 20, pp. 91-124.
- Nijland, H.A., E. van Kempen, J. Jabben and J.A. Annema (2001), *Geluidmaatregelen: kosten en baten*, RIVM report 715120005, Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven.
- Nolan, J.F. (1996), Determinants of productive efficiency in urban transit, *Logistics and Transportation Review*, vol. 32, pp. 319-342.
- Nolan, J.F., P.C. Ritchie and J.E. Rowcroft (2002), Identifying and measuring public policy goals: ISTEA and the U.S. bus transit industry, *Journal of Economic Behavior and Organization*, vol. 48, pp. 291-304.
- Nollet, C., B. Thiry and H. Tulkens (1988), Mesure de l'efficacité productive: application a la société de transports intercommunaux de Bruxelles, in: B. Thiry and H. Tulkens (eds.), *La Performance Economique des Sociétés Belges de Transport Urbains*, CIRIEC, Charleroi, pp. 137-170.
- Obeng, K. (1994), The economic cost of subsidy-induced technical inefficiency, *International Journal of Transport Economics*, vol. 21, pp. 3-20.
- OECD/ECMT (1995), *Urban Travel and Sustainable Development*, Organization for Economic Cooperation and Development/European Conference of Ministers of Transport, Paris



- Oertli, J. and D. Wassmer (1996), Rail noise control in Switzerland: legislation, environment, politics and finances, *Journal of Sound and Vibration*, vol. 193, pp. 403-406.
- Okeahialam, C.C. (1990), *An empirical study of cross-channel route choice*, Channel Tunnel Research Unit, University of Kent at Canterbury.
- Oum, T.H. and D.W. Gillen (1983), The structure of intercity travel demands in Canada: theory tests and empirical results, *Transportation Research B*, vol. 17, pp. 175-191.
- Oum, T.H., A. Zhang and Y. Zhang (1993), Inter-firm rivalry and firm-specific price elasticities in the deregulated airline markets, *Journal of Transport Economics and Policy*, vol. 27, pp. 171-192.
- Oum, T.H., D.W. Gillen and S.E. Noble (1986), Demand for fare classes and pricing in airline markets, *Logistics and Transportation Review*, vol. 22, pp. 195-222.
- Oum, T.H., W.G. Waters and J.S. Yong (1992), Concepts of price elasticities of transport demand and recent empirical estimates, *Journal of Transport Economics and Policy*, vol. 26, pp. 139-154.
- Oxley, L. (1994), Cointegration, causality, and Wagner's law: a test for Britain, 1870-1913, *Scottish Journal of Political Economy*, vol. 41, pp. 286-298.
- Peeters, A.L., R.G. de Jong, J.P. Kaper and J.C. Tukker (1982), *Hinder door spoorweggeluid in de woonomgeving*, Report D60, IMG-TNO, Delft
- Phillips, J.M. and E.P. Goss (1985), The effect of state and local taxes on economic development: A meta-analysis, *Southern Economic Journal*, vol. 62, pp. 320-333.
- Phlips, L. (1972), A dynamic version of the Linear Expenditure Model, *Review of Economics and Statistics*, vol. 54, pp. 450-458.
- Pindyck, R. (1979), Interfuel substitution and the industrial demand for energy: an international comparison, *Review of Economics and Statistics*, vol. 61, pp. 169-179.
- Quinet, E. (2004), A meta-Analysis of Western European external costs estimates, *Transportation Research D*, vol. 9, pp. 465-476.
- Rademaker, B.C., C.G.J. Schotten and B.A.M. Staatsen (1996), *Berekening van het aantal geluidgehinderden door railverkeer in Nederland*, RIVM rapport 715101004, Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven.
- Ramanathan, R. (1999), Short and long run elasticities of gasoline demand in India: an empirical analysis using cointegration techniques, *Energy Economics*, vol. 21, pp. 321-330.
- Ramsey, J.B., R. Rasche and B.T. Allen (1975), An analysis of the private and commercial demand for gasoline, *Review of Economics and Statistics*, vol. 57, pp. 502-507.
- Raudenbush, S.W. (2002), *Hierarchical linear models applications and data analysis methods*, Sage Publications, Thousand Oaks.

- Reggiani, A., P. Nijkamp and E. Sabella (1999), New advances in spatial network modeling, *European Journal of Operational Research*, vol. 128, pp. 385–401.
- Reza, A. and M. Spiro (1979), The demand for passenger car transport services and for gasoline, *Journal of Transport Economics and Policy*, vol. 13, pp. 304-319.
- Ricci, A. (2004), *Urban transport pricing*, ISIS Discussion Paper, Istituto di Studi per l'Integrazione dei Sistemi, Roma
- Rietveld P. and E. Verhoef (1998), Social feasibility of policies to reduce externalities in transport, in: K. Button and E. Verhoef (eds), *Road pricing, traffic congestion and the environment*, Edward Elgar, Cheltenham.
- Ripley, A. (2003), *The cars that ate London, Paris, Brussels, Amsterdam, Rome, Madrid, Vienna, Athens...*, Eco Dalle Citta, february 2003.
- Rodenburg, C.A., B. Ubbels and P. Nijkamp (2002), Policy scenarios for achieving sustainable transportation in Europe, *Transport Reviews*, vol. 22, pp. 449-472.
- Romilly, P., H. Song and X. Liu (1998), Modelling and forecasting car ownership in Britain, *Journal of Transport Economics and Policy*, vol. 32, pp. 165-185.
- Rosenthal, R. (1984), *Meta-analytical procedures for social research*, Applied Social Research Methods Series vol. 6., Sage Publications, Beverly-Hills.
- Rosenthal, R. and D.B. Rubin (1978), Interpersonal expectancy effects: the first 345 studies, *The Behavioral and Brain Sciences*, vol. 3, pp. 377-415.
- Rothengatter, W. (1993), Externalities of Transport, in J. Polak and A. Heertje (eds.), *European transport economics*, Blackwell, Oxford .
- Saad, M. (1983), Demand for air passenger travel in Australia, *Papers of the Eighth Australasian Transport Research Forum*, vol. 2, pp. 605-623.
- SAI (2001), *The Institute of Shipping Analysis*, Sweden
- Sakano, R., K. Obeng, K. and G. Azam (1997), Subsidies and inefficiency: stochastic frontier approach, *Contemporary Economic Policy*, vol. 15, pp. 113-127.
- Sakano, R. and K. Obeng (1995), Re-examination of inefficiencies in urban transit systems: a stochastic frontier approach, *Logistics and Transportation Review*, vol. 31, pp. 377-392.
- Samimi, R. (1995), Road transport energy demand in Australia: a cointegrated approach, *Energy Economics*, vol. 17, pp. 329-339.
- Schipper, Y. (1999), *Market structure and environmental costs in aviation. a welfare analysis of European air transport reform*, Doctorate Thesis, Tinbergen Institute Research Series 217, Amsterdam
- Schipper, Y., P. Nijkamp and P. Rietveld (2001), Aircraft noise valuation studies and meta-Analysis, *International Journal of Environmental Technology and Management*, vol. 1, pp. 317-320

- Schmidt, F.L. and J.E. Hunter (1977), Development of a general solution to the problem of validity generalization, *Journal of Applied Psychology*, vol. 62, pp. 529-540.
- Schwartz, J. (1994), Air pollution and daily mortality: A review and a meta-analysis, *Environmental Economics*, vol. 64, pp. 36-52.
- Small, K.A. (1992), *Urban transportation economics*, Harwood Academic Publishers, Chur.
- Smith, A.B. and J.N. Toms (1987), *Factors affecting demand for international travel to and from Australia*, Bureau of Transport Economics, Canberra.
- Smith, M.L. and G.V. Glass (1977), Meta-analysis of psychotherapy outcome studies, *American Psychologist*, vol. 32, pp. 752-760.
- Smith, M.L., G.V. Glass and T.I. Miller (1980), *The benefits of psychotherapy*, John Hopkins University Press, Baltimore.
- Smith, V.K. and J.-C. Huang (1995), Can markets value air quality? A meta-analysis of hedonic property value models, *Journal of Political Economy*, vol. 103, pp. 209-225.
- Smith, V.K. and K. Pattanayak (2002), Is meta-analysis a Noah's ark for non-market valuation, *Environmental and Resource Economics*, vol. 22, pp. 271-296.
- Smith, V.K. and L. Osborne (1996), Do contingent valuation estimates pass a "scope" test? A meta-analysis, *Journal of Environmental Economics and Management*, vol. 31, pp. 287-301.
- Stanley, T.D. (1998), New wine in old bottles: A meta-analysis of Ricardian equivalence, *Southern Economic Journal*, vol. 64, pp. 713-727.
- Stanley, T.D. (2005), Beyond publication bias, *Journal of Economic Surveys*, Vol. 19, pp. 309-346.
- Stanley, T.D. and S.B. Jarrell (2005), Meta-regression analysis: a quantitative method of literature surveys, *Journal of Economic Surveys*, vol. 19, pp. 299-308.
- Stigler, S.M. (1986), *The history of statistics: the measurement of uncertainty before 1900*, Harvard University Press, Cambridge.
- Stigler, S.M. (1996), Statistics and the question of standards, *Journal of Research of the National Institute of Standards and Technology*, vol. 101, pp. 779-789.
- Strand, J. and M. Vagnes (2001), The relationship between property values and railroad proximity: a study based on hedonic prices and real estate brokers' appraisals, *Transportation*, vol. 28, pp. 137-156.
- Straszheim, M.R. (1978), Airline demand functions in the North Atlantic and their pricing implications, *Journal of Transport Economics and Policy*, vol. 12, pp. 179-195.
- Summers, R. and A. Heston (1991), The Penn world table (Mark 5): an expanded set of international comparisons, 1950-1988, *The Quarterly Journal of Economics*, vol. 106, pp. 327-368.

- Sutton, A.J., K.R. Abrams, D.R. Jones, T.A. Sheldon and F. Song (2000), *Methods for meta-analysis in medical research*, Wiley, Chichester.
- Talley, W.K. and A. Schwarz-Miller (1988), The demand for air services provided by air passenger-cargo carriers in a deregulated environment, *International Journal of Transport Economics*, vol. 15, pp. 159-167.
- Taplin, J.H.E. (1980), A coherence approach to estimates of price elasticities in the vacation travel market, *Journal of Transport Economics and Policy*, vol. 14, pp. 19-35.
- Taplin, J.H.E. (1997), A generalised decomposition of travel-related demand elasticities into choice and generation components, *Journal of Transport Economics and Policy*, vol. 31, pp. 183-191.
- Thiry, B. and H. Tulkens (1992), Allowing for inefficiency in parametric estimation of production functions for urban transit firms, *Journal of Productivity Analysis*, vol. 3, pp. 45-65.
- Thissen, M. (1999), *A classification of empirical CGE modelling*, Research Report 99C02, Graduate school / Research Institute for Systems, Organisations and Management, University of Groningen.
- Thompson, A. and R. Caves (1992), The projected market share for a new small airport in the North of England, *Regional Studies*, vol. 27, pp. 137-147.
- Thompson, S.G., R.M. Turner and D.E. Warn (2001), Multilevel models for meta-analysis, and their application to absolute risk differences, *Statistical Methods in Medical Research*, vol. 10, pp. 375-392.
- Tishler, A. (1980), *The demand for cars and the price of gasoline: the user cost approach*, Working Paper, Foerder Institute for Economic Research, Tel Aviv University.
- Tishler, A. (1983), The demand for cars and gasoline: a simultaneous approach, *European Economic Review*, vol. 20, pp. 271-287.
- Tolley, R. (ed.) (2004), *Sustainable transport*, Woodhead Publishing Limited, Cambridge.
- Tone, K. and T. Sawada (1990), An efficiency analysis of public vs. private bus transportation enterprises, in: H. Bradley (ed.), *Operational research '90*, Pergamon, New York, pp.357-365.
- Tulkens, H. (1993), On FDH efficiency analysis: some methodological issues and applications to retail banking, courts and urban transit, *Journal of Productivity Analysis*, vol. 4, pp. 183-210.
- Tulkens, H. and P. Vanden Eeckaut (1995), Non-frontier measures of efficiency, progress and regress for time series data, *International Journal of Production Economics*, vol. 39, pp. 83-97.
- Tulkens, H. and P. Wunsch (1994), Les Performance économiques de la STIB au mois le mois, *Revue Suisse d'Économie Politique et de Statistique*, vol. 130, pp. 627-646.

- Tulkens, H., B. Thiry and A. Palm (1988), Mesure de l'efficacité productive: methodologies et applications aux sociétés de transports intercommunaux de Liege, Charleroi et Verviers, in: B. Thiry and H. Tulkens (eds.), *La Performance Economique des Sociétés Belges de Transport Urbains*, CIRIEC, Charleroi, pp. 81-136.
- Tyssen, L. (1982), *Consequenties van voorgenomen woningbouwplannen van het aanhouden van grenswaarden met betrekking tot spoorweglawaaï*, Report 4839-A, Stad en Landschap, Adviesbureau voor Ruimtelijke Ordening en Vormgeving B.V., Rotterdam.
- Uri, N. and S. Hassanein (1985), Testing for stability: motor gasoline demand and distillate fuel oil demand, *Energy Economics*, vol. 7, pp. 87-92.
- Bergh, J.C.J.M. van den, K.J. Button, P. Nijkamp and G.C. Pepping (1997). *Meta-analysis in environmental economics*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Verhoef, E.T. (1994), External effects and social costs of road transport, *Transportation Research A*, vol. 28, pp. 273-287.
- Vickrey, W. (1968), Pricing in urban and suburban transport, *American Economic Review*, vol. 53, pp. 452-465.
- Viton, P. (1986), The question of efficiency in urban bus transportation, *Journal of Regional Science*, vol. 26, pp. 499-513.
- Viton, P. (1992), Consolidations of scale and scope in urban transit, *Regional Science and Urban Economics*, vol. 22, pp. 25-49.
- Viton, P. (1997), Technical efficiency in multimode bus transit: a production frontier analysis, *Transportation Research B*, vol. 31, pp. 23-39.
- Viton, P. (1998), Changes in multimode bus transit efficiency, 1988-1992, *Transportation*, vol. 25, pp. 1-21.
- VTPI, *Online TDM Encyclopedia*, Victoria Transport Policy ([www.vtpi.org](http://www.vtpi.org)), 2002.
- Walters, A.A. (1961), The theory and measurement of private and social cost of highway congestion, *Econometrica*, vol. 29, pp. 676-697.
- Walters, A.A. (1968), *The economics of road user charges*, John Hoskins University Press, Baltimore.
- Walters, A.A. (1978), Airports - an economic survey, *Journal of Transport Economics and Policy*, vol. 12, pp. 125-160.
- Weinberger, M., H.G. Thomassen and R. Willeke (1991), *Kosten des Lärms in der Bundesrepublik Deutschland*, Erich Schmidt Verlag, Berlin.
- Wheatcroft, S. (1956), *The economics of European air transport*, Manchester University Press.
- Wheaton, W. (1982), The long-run structure of transportation and gasoline demand, *The Bell Journal of Economics*, vol. 13, pp. 439-454.

- White, H. (1980), A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, vol. 48, pp. 817-838.
- Wirl, F. (1991), Energy demand and consumer price expectations: an empirical investigation of the consequences from the recent oil price collapse, *Resources and Energy*, vol. 13, pp. 241-262.
- Wit, J. de and H. van Gent (2001), *Economie en Transport*, Lemma BV, Utrecht.
- Woodward, R.T. and Y.-S. Wui (2001), The economic value of wetland services: A meta-analysis, *Ecological Economics*, vol. 37, pp. 257-270.
- Wunsch, P. (1994), *Costing busses: back to basics*, FUSL (SMASH Cahier 9405), Brussels.
- Zhang, Y. and Bartels R. (1998), The effect of sample size on the mean efficiency in DEA: with an application to electricity distribution in Australia, Sweden and New Zealand, *Journal of Productivity Analysis*, vol. 9, pp. 187-204.



## Samenvatting (Summary in Dutch)

Verbeteringen in levensstandaarden en vervoerssystemen hebben in de 20ste eeuw wereldwijd geleid tot een spectaculaire toename van de vraag naar gemotoriseerd vervoer. Deze ontwikkelingen hebben bijgedragen aan regionale bereikbaarheid en economische groei maar hebben tevens geleid tot een toename van externe kosten als gevolg van onder andere luchtvervuiling, geluidsoverlast, verkeersongevallen en filevorming. Dit heeft bij nationale en internationale overheden geleid tot meer aandacht voor de duurzaamheid van vervoersbeleid. Belangrijke aandachtspunten bij een duurzaam vervoersbeleid zijn onder andere een nauwkeurige waardering van externe effecten, een effectief prijsbeleid en efficiënt werkende OV-systemen. In deze dissertatie richten wij ons op het toepassen van bestaande en nieuw ontwikkelde meta-analytische technieken met als doel verklaringen te vinden voor de empirische variatie in (i) de prijselasticiteit van de vraag naar luchtvaart; (ii) de prijselasticiteit van de vraag naar autobrandstof; (iii) de technische efficiency van het stedelijk openbaar vervoer; (iv) de economische eenheidskosten van geluidsoverlast van railverkeer. Hierbij richten we ons op het schatten van gemiddelde waarden voor deze indicatoren alsmede het identificeren van de invloed van studiekekenmerken en conditionerende factoren. Een veel voorkomend probleem bij meta-analyse in de economie is correlatie tussen observaties afkomstig uit dezelfde studie. Teneinde de toepassing van efficiënte schattingsmethoden te waarborgen, voeren we eerst een Monte Carlo-studie uit waarin we de schattingsprestaties van verschillende schattingstechnieken in de aanwezigheid van zulke correlatie binnen studies vergelijken.

## Methodologische Inzichten

In hoofdstuk 2 beschrijven we de achtergrond, de kenmerken en de voordelen van meta-analyse als onderzoeksmethode. Meta-analyse is gebaseerd op kwantitatieve samenvoeging van bestaande studie-uitkomsten met betrekking tot een bepaald onderzoeksobject, ook wel 'effect-size' genoemd. Het grotendeels niet-experimentele karakter van economisch onderzoek veroorzaakt relatief veel heterogeniteit tussen studies, hetgeen bij meta-analyse leidt tot systematische variatie in de 'effect-sizes'. Hierdoor lijkt voor ons onderzoek de toepassing van multivariate meta-analytische technieken, in combinatie met beschrijvende statistiek, een



geschikte benadering. Wij volgen deze aanpak in drie van de vier empirische studies in deze dissertatie, waarbij we door middel van beschrijvende statistiek de variatie in 'effect-sizes' onderzoeken en door middel van meta-regressieanalyse deze variatie verklaren. Meta-regressieanalyse stelt ons in staat om studieresultaten uit een grote verscheidenheid aan studies te combineren, door te corrigeren voor culturele, geografische, economische en institutionele factoren en studiekenmerken. Door het combineren van resultaten, verkleinen we het effect van 'sampling error', wat leidt tot nauwkeurigere schattingen van de 'effect-size'. In de toegepaste meta-analytische studies leverde deze aanpak gemiddelde waarden op voor de prijselasticiteit van de vraag naar luchtvaart, de prijselasticiteit van de vraag naar autobrandstof en de technische efficiency van het stedelijk openbaar vervoer. Metaregressie stelt ons verder in staat om de invloed van conditionerende variabelen op de 'effect-size' te analyseren. In de studie naar de vraag naar luchtvaart en autobrandstof levert dit belangrijke inzichten in het vraaggedrag van de consumenten op. In de studie naar efficiency van OV-organisaties leidt dit interessante conclusies ten aanzien van de potentiële voordelen van een gedereguleerde omgeving voor het stedelijk openbaar vervoer. De empirische resultaten en beleidsimplicaties die hierna worden besproken illustreren dat meta-analyse een waardevolle onderzoeksmethode is die tot inzichten kan leiden die op basis van primair onderzoek niet hadden kunnen worden blootgelegd.

Naast de genoemde voordelen bespreken we ook een aantal potentiële problemen bij het toepassen van meta-analyse: 'publication bias', heterogeniteit en 'multiple sampling'. 'Publication bias' is aanwezig wanneer de kans op publicatie van een studie wordt beïnvloed door bepaalde factoren zodat de steekproef van gepubliceerde studies niet representatief is voor de totale populatie. Er zijn methoden voor detectie en correctie van 'publication bias' ontwikkeld maar de toepasbaarheid daarvan in meta-analyses in de economie, gekenmerkt door 'multiple sampling' technieken en een grote mate van heterogeniteit tussen studies, is onzeker.

De heterogeniteit tussen economische onderzoeksresultaten kan leiden tot systematische en niet-systematische variatie in 'effect-sizes'. Voor systematische variatie kan worden gecorrigeerd door middel van meta-regressieanalyse. Niet-systematische variatie kan worden gecorrigeerd door het gebruik van schattingsmethoden zoals de 'mixed effects' schatter. Wanneer de dataset is gebaseerd op 'multiple sampling' technieken, geldt dat niet-systematische variatie in 'effect-sizes' in werkelijkheid kan worden veroorzaakt door statistische correlatie tussen observaties uit dezelfde studie. Wanneer hiervoor niet wordt gecorrigeerd leidt dit tot inefficiënte schattingen. Door middel van een Monte Carlo-studie

vergelijken we de schattingsprestaties van diverse schattingsmethoden waaronder OLS en een '2-level' schatter in de aanwezigheid van correlatie binnen studies. Schattingsmethoden die rekening houden met de afhankelijkheidsstructuur in de dataset blijken beter te presteren dan methoden die dat niet doen. Het 2-level model blijkt het hierbij het best te doen. Op basis van dit resultaat pleiten we bij toegepaste meta-analytische studies voor (i) het gebruik van een statistische test gebaseerd op het '2-level' model om te testen op correlatie binnen studies en, indien aanwezig, (ii) het gebruik van het '2-level' model als schattingsmethode. Wij volgen deze aanpak in de toegepaste meta-analytische studies in deze dissertatie. Bij de studies naar de prijsgevoeligheid van luchtvaart en de technische efficiency van openbaar vervoer wezen statische testen op de aanwezigheid van correlatie binnen studies. In deze studies gebruiken wij dan ook het '2-level' model als schattingsmethode.

In de studie naar de prijsgevoeligheid van de vraag naar autobrandstof ontwikkelen we een econometrische methode, gebaseerd op een systeem van vergelijkingen, waarbij tegelijkertijd meerdere metaregressievergelijkingen kunnen worden geschat, door gebruik te maken van de lineaire relaties tussen de verschillende 'effect-sizes'. Deze methode heeft twee voordelen ten opzichte van gebruikelijke meta-analytische technieken. Ten eerste is de hoeveelheid informatie die gebruikt kan worden in een dergelijke analyse groter, hetgeen leidt tot nauwkeurigere schattingen. Ten tweede is het mogelijk om de 'effect-size' uit te splitsen in een aantal deeleffecten waarbij tevens de invloed van conditionerende variabelen op de 'effect-size' nader kan worden geanalyseerd door te kijken naar de invloed per deeleffect. Wanneer aan bepaalde voorwaarden wat betreft het aantal observaties wordt voldaan, lijkt deze methode een nuttige schattingstechniek voor meta-analyses waarvan de 'effect-size' kan worden gesplitst in een aantal deeleffecten die op zich zelf interessant zijn om te onderzoeken.

## **Conclusies met Betrekking tot Empirische Resultaten**

In Hoofdstuk 4 beschrijven we een meta-analytische studie waarin we de empirische variatie in de prijselasticiteit van de vraag naar luchtvaart onderzoeken. Gebaseerd op een verzameling van 204 observaties vinden we een gemiddelde prijselasticiteit van -1.15, hetgeen betekent dat prijswijzigingen resulteren in een meer dan proportionele wijziging in de vraag. Door middel van meta-regressieanalyse onderzoekt deze studie tevens de invloed van diverse conditionerende factoren op de prijselasticiteit. De belangrijkste resultaten zijn als volgt. Elasticiteiten zijn op lange termijn lager dan op korte termijn, wat wijst op de

significantie van een langere responstijd. Reizigers zijn in de loop van de tijd prijsgevoeliger geworden, wellicht door toenemende afhankelijkheid als gevolg van de groei in de luchtvaartsector. 'Business class' reizigers zijn minder prijsgevoelig dan 'economy class' reizigers, wat onder andere verklaard kan worden door verschillen in substitutiemogelijkheden en tijdswaardering. Ondanks regionale verschillen in substitutiemogelijkheden vinden we geen verschillen in prijsgevoeligheid tussen regio's. Studies die niet voor inkomen corrigeren vinden gemiddeld hogere prijsgevoeligheid. Dit resultaat geeft aan dat in nieuwe casestudies hiervoor gecorrigeerd zou moeten worden.

In de meta-analytische studie beschreven in Hoofdstuk 5, richten we ons op de prijselasticiteit van de vraag naar autobrandstof. De prijselasticiteit van de totale vraag naar brandstof kan worden gezien als de som van drie deelelasticiteiten met betrekking tot brandstofefficiency, autogebruik en autobezit. In deze studie schatten we gemiddelde waarden voor de prijselasticiteit van de vraag naar autobrandstof en voor elk van de deelelasticiteiten, waarbij we de lineaire relatie tussen de elasticiteiten als schattingsrestrictie hanteren.

De gemiddelde prijselasticiteit van de vraag naar autobrandstof blijkt met een waarde van -0.53 inelastisch te zijn. Uit de resultaten met betrekking tot de deelelasticiteiten blijkt dat de invloed van een prijswijziging op de vraag voor een groot deel kan worden verklaard door wijzigingen in de brandstofefficiency (0.22) en het autobezit (-0.22) en in minder mate door wijzigingen in het autogebruik (-0.10). Vervolgens passen we meta-regressieanalyse toe om de invloed van conditionerende factoren op de prijselasticiteiten te identificeren. Hierbij schatten we meerdere metaregressievergelijkingen tegelijkertijd waarbij we de lineaire relatie tussen de prijselasticiteiten als schattingsrestrictie hanteren. De belangrijkste resultaten zijn als volgt. De prijsgevoeligheid is lager in de VS, Canada en Australië dan in de rest van de wereld, hetgeen voornamelijk wordt bepaald door een lagere prijsgevoeligheid met betrekking tot het autobezit. De hoge afhankelijkheid van het vervoer per auto in deze landen kan worden verklaard door de combinatie van hoge inkomens, weinig ontwikkeld openbaar vervoer en relatief lage brandstofprijzen. De prijsgevoeligheid is hoger op de lange termijn. Dit blijkt voornamelijk te worden veroorzaakt door langetermijnaanpassingen in het aantal kilometers per voertuig. Verder blijkt dat autobezitters in de loop van de tijd prijsgevoeliger zijn geworden, hetgeen verklaard zou kunnen worden vanuit een toename van het deel van het inkomen dat wordt uitgegeven aan autobrandstof. Andere studiekenmerken die van invloed zijn op de prijselasticiteit zijn het gebruik van een dynamische dan wel statische modelspecificatie en het type database (tijdreeksen, panel data, etc.) dat gebruikt wordt.

In Hoofdstuk 6 onderzoeken we met behulp van meta-analyse de empirische variatie in de gemiddelde technische efficiency van stedelijke OV-organisaties. Op basis van een dataset van 93 observaties vinden we een gemiddelde waarde voor technische efficiency van 0.825. Vervolgens gebruiken we meta-regressieanalyse om de invloed van conditionerende factoren te identificeren. De relaties die we vinden tussen de technische efficiency enerzijds en het inkomensniveau, het aandeel van overheidsuitgaven in het BNP, diverse geografische dummyvariabelen en een trendvariabele anderzijds suggereren dat een gedereguleerde omgeving bevorderend is voor de efficiency van OV-organisaties. Andere studiekenmerken die de technische efficiency beïnvloeden zijn: de steekproefgrootte, het type database (tijdreeksen, panel data, etc.), de veronderstellingen waarop de 'efficiency frontier' gebaseerd is, het aantal inputs waarop de schatting gebaseerd is en de outputmaatstaf die hierbij gehanteerd is.

In hoofdstuk 7 beschrijven we een review studie naar de eenheidskosten van geluidsoverlast van railverkeer. In deze studie wordt een conceptueel model beschreven waarin railverkeerkenmerken, geluidsemissiewaarden, geluidsimmissiewaarden en geluidsoverlast een causale keten vormen. Geluidsoverlast leidt uiteindelijk tot sociale en economische kosten, zoals waardedalingen van onroerend goed. De economische kosten van geluidsoverlast blijken niet alleen afhankelijk te zijn van het geluidsniveau maar tevens van kenmerken zoals het soort geluid, de frequentie van de geluidsemissie, de temporele verdeling van de geluidsemissie en persoonskenmerken zoals: gebruik van het spoorvervoer, gewinning, houding en activiteitenpatronen. Beleidsmaatregelen gericht op het reduceren van de sociale en economische kosten zijn binnen dit conceptuele model op diverse plaatsen van belang. Beleidsmaatregelen kunnen worden onderverdeeld in regulering en preventiemaatregelen. Strengere regulering met betrekking tot geluidslimieten blijkt te leiden tot hogere totale kosten maar kan tot lagere sociale kosten per hoofd van de bevolking leiden. De economische haalbaarheid van overheidsmaatregelen wordt veelal onderzocht door middel van kosten-batenanalyses. Onderzoeksmethoden zijn echter divers en vaak ad hoc, zodat resultaten van verschillende case studies moeilijk met elkaar te vergelijken zijn. De heterogeniteit tussen casestudies heeft geleid tot de toepassing van een breed scala aan waarderingsmethoden. Een combinatie van waarderingsmethoden is veelal de meest optimale benadering voor een realistische waardering van de externe kosten.

## Beleidsimplicaties

Op 12 september 2001 presenteerde de Europese Commissie het zogenaamde 'Witboek' getiteld "Het Europese vervoersbeleid tot het jaar 2010: tijd om te kiezen" met betrekking tot het toekomstige gemeenschappelijke vervoersbeleid. De belangrijkste doelstellingen van het gemeenschappelijke vervoerbeleid waren: (i) het hanteren van heffingsystemen die rekening houden met de daadwerkelijke kosten van vervoer, inclusief externe kosten zoals milieuschade, filevorming en verkeersongevallen; (ii) het verhogen van de efficiency en de veiligheid van vervoerssystemen; (iii) het verschuiven van de balans tussen vervoersmodaliteiten per 2010 door het revitaliseren van spoorwegen en het stimuleren van het vervoer per water.

Prijstoeslagen vormen een belangrijk economisch beleidsinstrument. Als gevolg van de toetreding van 'low cost carriers' is het aandeel van prijsheffingen in de gemiddelde totale kosten van luchtvaartvervoer toegenomen. Het potentiële effect van prijsstoeslagen in de luchtvaartsector lijkt daarom te zijn toegenomen. De Europese Commissie pleit voor een uiteindelijke afschaffing door lidstaten van de ontheffingen van brandstoftoeslag die traditioneel toegepast worden in de luchtvaartsector. Momenteel is een brandstoftoeslag op vluchten binnen lidstaten toegestaan, maar voor internationale vluchten is de toepassing ervan vaak nog onmogelijk. Verder benadrukt de EU het belang van de introductie van economische en regulerende marktinstrumenten om de concurrentiepositie te verbeteren van organisaties die kiezen voor state-of-the-art technologieën en milieuvriendelijke bedrijfsvoering. De EC beschouwt een systeem van gelimiteerde, verhandelbare vervuilingsrechten als het meest effectieve economische instrument om luchtvaartmaatschappijen te bewegen tot milieuvriendelijkere bedrijfsvoering.

Voor een nauwkeurige schatting van de effecten van prijsbeleid, heeft de overheid informatie nodig omtrent de prijsgevoeligheid van reizigers. In de studie naar de prijsgevoeligheid van de vraag naar luchtvaart vinden we een gemiddelde prijselasticiteit van -1.15, hetgeen wijst op aanzienlijke prijsgevoeligheid onder luchtvaartreizigers; een prijsverhoging leidt tot een meer dan proportionele afname van de vraag. De resultaten van de meta-regressieanalyse wijzen erop dat 'business class' reizigers minder prijsgevoelig zijn dan 'economy class' reizigers. Dit geeft de luchtvaartmaatschappijen de mogelijkheid om de extra kosten tengevolge van een prijsheffing gedeeltelijk door te berekenen aan de passagiers, waarbij de toeslag voor 'business class' reizigers hoger is dan de toeslag voor 'economy class'

reizigers. Een effectief prijsbeleid dient hier rekening mee te houden. Over het algemeen wijzen de studieresultaten erop dat prijshellingen een effectief instrument zijn om luchtvaartmaatschappijen tot milieuvriendelijkere beleidsvoering te bewegen. De extra inkomsten als gevolg van de prijshellingen kunnen hierbij worden gebruikt ter financiering van subsidies voor het gebruik van milieuvriendelijke technologieën in de luchtvaart. Het huidige EU-beleid met betrekking tot brandstofhellingen voor binnenlandse vluchten en de EU beleidsvoorstellen met betrekking tot de verdere afbouw van ontheffingen van brandstoftoeslag en de stimulatie van milieuvriendelijke technologieën en beleidsvoering kunnen de ontwikkeling en toepassing van dergelijke heffingsystemen ondersteunen.

Met betrekking tot het belasten van het wegverkeer wordt binnen de EU een breed scala aan fiscale instrumenten toegepast. De discussie met betrekking tot economische beleidsinstrumenten richt zich met name op het verbeteren van heffingsystemen met als doel de gebruiker te confronteren met de daadwerkelijke kosten. Recente beleidsontwikkelingen wekken de indruk dat het vervoersbeleid in Europa zich in toenemende mate zal baseren op variabilisatie, hetgeen zal leiden tot een verschuiving van bezitsgeoriënteerde instrumenten zoals wegenbelasting naar gebruiksgeoriënteerde instrumenten zoals brandstofhellingen en rekeningrijden. Recent Europees beleid op het gebied van brandstofhellingen richt zich met name op het harmoniseren van toeslagen in lidstaten en op de promotie van het gebruik van biobrandstoffen of andere hernieuwbare brandstoffen voor vervoer; EU-richtlijnen schrijven minimum toeslagen voor onder andere motorbrandstof voor, waarbij het lidstaten toegestaan is volledige of gedeeltelijke ontheffingen of verlagingen toe te passen voor onder andere biobrandstoffen. Stimulering van brandstofefficiency beperkt zich in de Europese Unie tot regulering met betrekking tot emissies en consumenteninformatie. In de VS gelden sinds 1975 zogenaamde 'CAFE-normen' voor brandstofefficiency van personenauto's; autofabrikanten die de normen niet halen worden geconfronteerd met financiële sancties. Er is veel discussie met betrekking tot de effectiviteit van dit beleid; voorstanders zien de CAFE-normen als een bewezen manier om de vraag naar brandstof te verlagen terwijl tegenstanders het beleid bekritiseren als kostbaar en omslachtig en brandstoftoeslagen als een effectiever instrument zien.

De resultaten van de studie besproken in hoofdstuk 5 lieten zien dat de vraag naar autobrandstof, met een gemiddelde prijselasticiteit van -0.53, niet erg prijsgevoelig is. Dit resultaat steunt het standpunt dat beleid gebaseerd op efficiëncynormen zoals CAFE wellicht een effectievere manier is om de vraag naar brandstof te verminderen. De resultaten laten verder zien dat een prijswijziging een sterkere invloed heeft op het niveau van

brandstofefficiency en het autobezit dan op het autogebruik. De relatief lage prijsgevoeligheid van het autogebruik, vergeleken met dat van autobezit, wijst erop dat in de discussie met betrekking tot variabilisatie de invloed van een brandstoftoeslag op het autobezit niet genegeerd kan worden. Al met al blijken brandstofheffingen niet erg effectief te zijn in het verminderen van de externe kosten van wegverkeer en lijken er aanvullende instrumenten nodig te zijn. Registratieheffingen en brandstoftoeslagen zijn van kracht in de meeste lidstaten maar volgens analisten is de huidige structuur en het niveau van deze heffingen niet bevorderend voor de verbetering van brandstofefficiency. De EU zou zich kunnen richten op de mogelijkheden om 'slimme' heffingssystemen te promoten, die het gebruik van brandstofefficiënte technologieën stimuleren door middel van belastingonthefingen en subsidiering. Hierbij kan men denken aan innovatieve belastingregelingen zoals het in Canada en Oostenrijk geïntroduceerde 'feebate' concept, dat gebaseerd is op heffingen voor voertuigen met hoge brandstofconsumptie en belastingkortingen voor brandstofefficiënte voertuigen.

In West-Europa is concurrentie in de OV-sector gedurende enkele decennia als gevolg van overheidsregulering grotendeels uitgesloten. Deze situatie is in de afgelopen tien jaar sterk gewijzigd. De meeste lidstaten hebben met betrekking tot (een deel van) de OV-sector een concurrentie-element in hun wetgeving of administratieve praktijk geïntroduceerd, meestal gebaseerd op regelmatige vernieuwing van exclusieve rechten, in plaats van op vrije toegang tot de markt. EU-wetgeving vereist momenteel alleen in de lucht- en zeevaartsector openbare aanbesteding van openbare dienstcontracten. Met betrekking tot het vervoer over land beperkt de wetgeving zich tot het opleggen van rekenmethoden voor de subsidiering van de uitvoer van openbare diensten. Recent heeft de EU een voorstel ter consultatie voorgelegd, gebaseerd op (i) contractuele regelingen tussen de betreffende autoriteiten en organisaties verantwoordelijk voor het leveren van diensten waarvoor financiële compensatie of exclusieve rechten zijn toegewezen; (ii) de noodzaak van periodieke beoordeling van contractuele afspraken; (iii) openbare aanbesteding van contracten voor de uitvoer van openbare diensten, behalve in bepaalde gevallen waar directe toekenning mogelijk is.

De resultaten van de meta-analyse beschreven in Hoofdstuk 6 suggereren dat een gedereguleerde omgeving bevorderend is voor de efficiency van OV-organisaties. Het voorstel van de EU, dat momenteel wordt besproken in het Europese parlement, is een stap in de richting van een verdere afname van de rol van de overheid. Hoewel de voordelen van de principes waarop het voorstel is gebaseerd worden ondersteund door de resultaten van de meta-analyse, zal het voorstel moeten worden ondersteund door onderzoek waarbij ook

aandacht wordt besteed aan de (negatieve) invloed van deregulering op aspecten als dienstfrequentie, regionale dekking, ritprijs en kwaliteit.

Europees beleid op het gebied van geluidsoverlast van railverkeer bestaat uit wetgeving met betrekking tot de interoperabiliteit van het trans-Europese railnetwerk, waarin eisen worden gesteld aan de technische specificatie van rijdend materieel en de railinfrastructuur. Nationaal overheidsbeleid is veelal gebaseerd op kosten-batenstudies waarbij de kosten van geluidsoverlast en preventiemaatregelen tegen elkaar afgewogen worden. Deze studies gebruiken vaak kostenindices van geluid in de vorm van NDSI-waarden afkomstig uit 'hedonic pricing' studies naar geluidsoverlast van weg- en luchtvaartverkeer. Uit de review studie naar geluidsoverlast van railverkeer blijkt dat de economische kosten niet alleen afhankelijk zijn van het geluidsniveau maar ook van diverse andere factoren, hetgeen de 'value transfer' van NDSI-waarden tussen modaliteiten compliceert en bij kosten-batenstudies kan leiden tot onder- of overwaardering van de daadwerkelijke kosten en dus tot suboptimale overheidsbeslissingen. Idealiter zou overheidsbeleid gebaseerd moeten zijn op NDSI-waarden van railverkeer, waarbij gecorrigeerd kan worden voor case-specifieke kenmerken en conditionerende factoren. Teneinde zulke NDSI-waarden te verkrijgen is meer primair onderzoek nodig.

## **Aanbevelingen voor Vervolgonderzoek**

Zowel het toegepaste als het methodologische onderzoek leidt tot een aantal interessante nieuwe onderzoeksvragen. In het algemeen geldt dat de toepassing van meta-analyse de behoefte aan primair onderzoek niet wegneemt. Meta-analyse kan slechts verricht worden op basis van een voldoende groot aantal primaire studies, zoals blijkt uit de studie naar geluidsoverlast van railverkeer. Het feit dat kosten-batenanalyses naar railverkeerbeleid veelal gebruik maken van NDSI-waarden afkomstig uit hedonic pricing studies naar geluidsoverlast van wegverkeer en luchtvaart, is een duidelijke indicatie dat er meer primair onderzoek nodig is naar NDSI-waarden van railverkeer. Hoewel de literatuur op het gebied van prijselasticiteiten van de vraag naar autobrandstof uitgebreid is, is het aantal observaties van elasticiteiten van autogebruik, autobezit en brandstofefficiëntie beperkt. Meer onderzoek naar deze elasticiteiten is nodig om de invloed van prijsheffingen nauwkeuriger te kunnen schatten. De luchtvaartsector blijft zich snel ontwikkelen. Regelmatig verschijnt er nieuws over faillissementen, samenwerkingsverbanden en toetredingen van 'low cost carriers'. Het



zou daarom nuttig kunnen zijn om de meta-analyse van prijselasticiteiten in de luchtvaart opnieuw te kunnen schatten op basis van nieuwe observaties en de resultaten te vergelijken. Vanwege het toenemende belang van 'low cost carriers', zou het interessant zijn om hierbij het absolute prijsniveau als verklarende variabele op te nemen. De resultaten van de studie naar de efficiency van het stedelijke OV-vervoer impliceren een negatieve relatie tussen de mate van overheidsregulering en technische efficiency. Deze conclusie zou nader kunnen worden onderzocht in een studie waarin een zo groot mogelijke verzameling van observaties uit zoveel mogelijk landen bijeengebracht wordt. Informatie met betrekking tot overheidssubsidie van het openbaar vervoer zou, indien beschikbaar, als verklarende variabele moeten worden toegevoegd. Een dergelijke studie heeft als voordeel dat alle observaties met een en dezelfde 'efficiency frontier' kunnen worden vergeleken. Bovendien geldt dat heterogeniteit veroorzaakt door verschillen in studiekenmerken afwezig is. Overigens geldt met betrekking tot elk van de in deze these toegepaste meta-analytische studies dat het interessant zou zijn om als vervolgonderzoek een 'state-of-the-art' primaire studie uit te voeren, rekening houdend met de inzichten die de meta-analyse opgeleverd heeft. Wanneer de resultaten van deze studies vergelijkbaar zijn met de op basis van de uitkomsten van de meta-analyses voorspelde resultaten, is dit een indicatie dat de meta-analytische modellen geschikt zijn voor value transfer en geconditioneerde voorspellingen.

Het ontwerp van de Monte Carlo studie beschreven in hoofdstuk 3 is gebaseerd op een aantal veronderstellingen ten aanzien van homoskedasticiteit, de correlatiestructuur en de correlatiegraad. Vanwege verschillen in steekproefgrootte tussen de onderliggende studies, lijkt de veronderstelling van homoskedasticiteit niet helemaal realistisch. Het zou daarom interessant zijn te onderzoeken hoe de resultaten zouden veranderen wanneer we uitgaan van heteroskedasticiteit binnen de dataset. Ook de veronderstelling dat de mate van correlatie binnen studies gelijk is voor alle studies is niet helemaal realistisch; voor heterogene studies zou men een lagere mate van correlatie binnen studies verwachten dan voor homogene studies. Men zou hier in de onderzoeksopzet rekening mee kunnen houden door per studie verschillende niveaus van correlatie te hanteren, eventueel gebaseerd op een indicator die de mate van heterogeniteit binnen de studie weergeeft. De studie is verder gebaseerd op het idee dat 'multiple sampling' statistische correlatie veroorzaakt tussen observaties afkomstig uit dezelfde studie, omdat deze tot stand zijn gekomen onder soortgelijke omstandigheden. Er zou echter ook statistische correlatie kunnen zijn tussen observaties afkomstig uit verschillende studies met soortgelijke kenmerken. Hier zou men rekening mee kunnen houden door een afhankelijkheidsstructuur te genereren waarbij de mate van correlatie tussen

observaties bijvoorbeeld gebaseerd is op het aantal studiekenmerken dat zij gemeen hebben. Ten slotte baseerden we ons op de veronderstelling dat de afhankelijkheid binnen studies in dezelfde mate aanwezig is in het geobserveerde en het niet-geobserveerde deel van de variatie. Hoewel deze veronderstelling gebaseerd is op realistische verwachtingen ten aanzien van de meta-analytische praktijk, zou het interessant zijn om de situatie te onderzoeken waarbij afhankelijkheid zich slechts voordoet in het niet-geobserveerde deel van de variatie.



The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

327. M.W. VAN GELDEREN, *Opportunity entry performance. Studies of entrepreneurship and small business.*
328. W. VAN WINDEN, *Essays on urban ICT policies.*
329. G.J. KULA, *Optimal retirement decision.*
330. R.J. IMESON, *Economic analysis and modeling of fisheries management in complex marine ecosystems.*
331. M. DEKKER, *Risk, resettlement and relations: Social security in rural Zimbabwe.*
332. A. MULATU, *Relative stringency of environmental regulation and international competitiveness.*
333. C.M. VAN VEELEN, *Survival of the fair: Modelling the evolution of altruism, fairness and morality.*
334. R. PHISALAPHONG, *The impact of economic integration programs on inward foreign direct investment.*
335. A.H. NÖTEBERG, *The medium matters: The impact of electronic communication media and evidence strength on belief revision during auditor-client inquiry.*
336. M. MASTROGIACOMO, *Retirement, expectations and realizations. Essays on the Netherlands and Italy.*
337. E. KENJOH, *Balancing work and family life in Japan and four European countries: Econometric analyses on mothers' employment and timing of maternity.*
338. A.H. BRUMMANS, *Adoption and diffusion of EDI in multilateral networks of organizations.*
339. K. STAAL, *Voting, public goods and violence.*
340. R.H.J. MOSCH, *The economic effects of trust. Theory and empirical evidence.*
341. F. ESCHENBACH, *The impact of banks and asset markets on economic growth and fiscal stability.*
342. D. LI, *On extreme value approximation to tails of distribution functions.*
343. S. VAN DER HOOG, *Micro-economic disequilibrium dynamics.*
344. B. BRYIS, *Tax-arbitrage in the Netherlands evaluation of the capital income tax reform of January 1, 2001.*
345. V. PRUZHANSKY, *Topics in game theory.*
346. P.D.M.L. CARDOSO, *The future of old-age pensions: Its implosion and*

*explosion.*

- 347. C.J.H. BOSSINK, *To go or not to go...? International relocation willingness of dual-career couples.*
- 348. R.D. VAN OEST, *Essays on quantitative marketing models and Monte Carlo integration methods.*
- 349. H.A. ROJAS-ROMAGOSA, *Essays on trade and equity.*
- 350. A.J. VAN STEL, *Entrepreneurship and economic growth: Some empirical studies.*
- 351. R. ANGLINGKUSUMO, *Preparatory studies for inflation targeting in post crisis Indonesia.*
- 352. A. GALEOTTI, *On social and economic networks.*
- 353. Y.C. CHEUNG, *Essays on European bond markets.*
- 354. A. ULE, *Exclusion and cooperation in networks.*
- 355. I.S. SCHINDELE, *Three essays on venture capital contracting.*
- 356. C.M. VAN DER HEIDE, *An economic analysis of nature policy.*
- 357. Y. HU, *Essays on labour economics: Empirical studies on wage differentials across categories of working hours, employment contracts, gender and cohorts.*
- 358. S. LONGHI, *Open regional labour markets and socio-economic developments: Studies on adjustment and spatial interaction.*
- 359. K.J. BENIERS, *The quality of political decision making: Information and motivation.*
- 360. R.J.A. LAEVEN, *Essays on risk measures and stochastic dependence: With applications to insurance and finance.*
- 361. N. VAN HOREN, *Economic effects of financial integration for developing countries.*
- 362. J.J.A. KAMPHORST, *Networks and learning.*
- 363. E. PORRAS MUSALEM, *Inventory theory in practice: Joint replenishments and spare parts control.*
- 364. M. ABREU, *Spatial determinants of economic growth and technology diffusion.*
- 365. S.M. BAJDECHI-RAITA, *The risk of investment in human capital.*
- 366. A.P.C. VAN DER PLOEG, *Stochastic volatility and the pricing of financial derivatives.*
- 367. R. VAN DER KRUK, *Hedonic valuation of Dutch Wetlands.*
- 368. P. WRASAI, *Agency problems in political decision making.*
- 369. B.K. BIERUT, *Essays on the making and implementation of monetary policy decisions.*
- 370. E. REUBEN, *Fairness in the lab: The effects of norm enforcement in economic decisions.*
- 371. G.J.M. LINDERS, *Intangible barriers to trade: The impact of institutions, culture, and distance on patterns of trade.*
- 372. A. HOPFENSITZ, *The role of affect in reciprocity and risk taking: Experimental studies of economic behavior.*
- 373. R.A. SPARROW, *Health, education and economic crisis: Protecting the poor in*

*Indonesia.*

- 374. M.J. KOETSE, *Determinants of investment behaviour: Methods and applications of meta-analysis.*
- 375. G. MÜLLER, *On the role of personality traits and social skills in adult economic attainment.*
- 376. E.H.B. FEIJEN, *The influence of powerful firms on financial markets.*
- 377. J.W. GROSSER, *Voting in the laboratory.*